GENETIC PSYCHOLOGY MONOGRAPHS

Child Behavior, Animal Behavior, and Comparative Psychology

EDITED BY
CARL MURCHISON

VOLUME XVI
1934

Issued monthly by the
CLARK UNIVERSITY PRESS
Worcester, Massachusetts
U. S. A.

TABLE OF CONTENTS

PSYCHOLOGIES OF 1925

By M. Bentley, K. Dunlap, W. S. Hunter, K. Koffka, W. Köhler, W. McDougall, M. Prince, J. B. Watson, and R. S. Woodworth. *Edited by* Carl Murchison.

Pp. xiii+412. $6.00 postpaid

PSYCHOLOGIES OF 1930

By A. Adler, M. Bentley, E. G. Boring, G. S. Brett, H. Carr, J. Dewey K. Dunlap, J. C. Flugel, W. S. Hunter, P. Janet, T. L. Kelley, K. Koffka, W. Köhler, K. N. Kornilov, W. McDougall, J. P. Nafe, I. P. Pavlov, F. Sander, A. L. Schniermann, C. Spearman, L. T. Troland, M. F. Washburn, A. P. Weiss, and R. S. Woodworth. *Edited by* Carl Murchison.

Pp. xiv+497. $6.00 postpaid

CRIMINAL INTELLIGENCE (1926)

By Carl Murchison, *Professor of Psychology and Director of the Psychological Laboratories in Clark University.*

Pp. 291. Out of print

THE CASE FOR AND AGAINST PSYCHICAL BELIEF (1927)

By Sir O. Lodge, Sir A. C. Doyle, F. B. Bond, L. R. G. Crandon, M. Austin, M. Deland, W. McDougall, H. Driesch, W. F. Prince, F. C. S. Schiller, J. E. Coover, G. Murphy, J. Jastrow, and H. Houdini. *Edited by* Carl Murchison.

Pp. 365. $3.75 postpaid

FEELINGS AND EMOTIONS: THE WITTENBERG SYMPOSIUM (1928)

By A. Adler, F. Aveling, V. M. Bekhterev, M. Bentley, G. C. Brett, K. Bühler, W. B. Cannon, H. A. Carr, E. Claparède, K. Dunlap, R. H. Gault, D. W. Gruehn, L. B. Hoisington, D. T. Howard, E. Jaensch, P. Janet, J. Jastrow, C. Jørgensen, D. Katz, F. Kiesow, F. Krueger, H. S. Langfeld, W. McDougall, H. Piéron, W. B. Pillsbury, M. Prince, C. E. Seashore, C. E. Spearman, W. Stern, G. M. Stratton, J. S. Terry, M. F. Washburn, A. P. Weiss, and R. S. Woodworth. *Edited by* M. L. Reymert.

Pp. xvi+454. $6.00 postpaid

SOCIAL PSYCHOLOGY: THE PSYCHOLOGY OF POLITICAL DOMINATION (1929)

By Carl Murchison, *Professor of Psychology and Director of the Psychological Laboratories in Clark University.*

Pp. x+210. $3.50 postpaid

THE COMMON SENSE OF DREAMS (1929)

By Henry J. Watt, *Late Lecturer in Psychology in the University of Glasgow, and Late Consulting Psychologist to the Glasgow Royal Asylum. Author of "The Psychology of Sound."*

Pp. xvi+212. $3.00 postpaid

THE FOUNDATIONS OF EXPERIMENTAL PSYCHOLOGY (1929)

By H. Banister, P. Bard, W. B. Cannon, W. J. Crozier, A. Forbes, S. I. Franz, F. N. Freeman, A. Gesell, H. Hartridge, S. Hecht, J. O. Holsopple, W. S. Hunter, T. L. Kelley, C. Landis, K. S. Lashley, M. A. May, T. H. Morgan, J. P. Nafe, G. H. Parker, R. Pintner, E. Stern, L. T. Troland, and C. Wissler. *Edited by* Carl Murchison.

Pp. x+907. $6.00 postpaid

CLARK UNIVERSITY PRESS
Worcester, Massachusetts
U. S. A.

LONDON: HUMPHREY MILFORD: OXFORD UNIVERSITY PRESS

$7.00 per volume
Single numbers $2.00

MONTHLY
Two volumes per year

July, 1934
Volume XVI, No. 1

GENETIC PSYCHOLOGY MONOGRAPHS

Child Behavior, Animal Behavior, and Comparative Psychology

EDITED BY
CARL MURCHISON

Worcester, Massachusetts

Entered as second-class matter December 1, 1925, at the post-office at Worcester, Mass., under Act of March 3, 1879

Journal of General Psychology

EDITED BY

CARL MURCHISON, *Clark University*

The Pedagogical Seminary and

Journal of Genetic Psychology

EDITED BY

CARL MURCHISON, *Clark University*

CLARK UNIVERSITY PRESS
Worcester, Massachusetts
U. S. A.

$7.00 per volume | MONTHLY | July, 1934
Single numbers $2.00 | Two volumes per year | Volume XVI, No. 1

GENETIC PSYCHOLOGY MONOGRAPHS

Child Behavior, Animal Behavior, and Comparative Psychology

A STATISTICAL STUDY OF RATINGS ON THE CALIFORNIA BEHAVIOR INVENTORY FOR NURSERY-SCHOOL CHILDREN*

From the Institute of Child Welfare, University of California

By

HERBERT S. CONRAD

*Accepted for publication by Harold E. Jones of the Editorial Board and received in the Editorial Office, May 25, 1933.

Worcester, Massachusetts

Entered as second-class matter December 1, 1925, at the post-office at Worcester, Mass., under Act of March 3, 1879.

ACKNOWLEDGMENT

The present study from the Institute of Child Welfare is based upon the ratings of traits from the California Behavior Inventory for Nursery School Children. The study could not have been undertaken successfully without the cooperation of many individuals. We are especially thankful to the three nursery-school teachers—Miss Lucille Allen, Mrs. Gladys Ludwig, and Miss Elsie Lewis—who assisted in the preparation of the Inventory and did the actual work of rating. To Dr. Herbert R. Stolz and Dr. Mary Cover Jones we are indebted for administrative assistance; to Miss Lina Hutson, for faithful and accurate aid in the statistical analysis; and to Dr. Harold E. Jones, Director of Research, for his interest and critical help in the project from its earliest stages to the final publication of the manuscript.

HERBERT S. CONRAD

INSTITUTE OF CHILD WELFARE
UNIVERSITY OF CALIFORNIA
BERKELEY, CALIFORNIA

CONTENTS

I

INTRODUCTION

A. The Data

In the spring of 1930, the three nursery-school teachers of the Institute of Child Welfare rated each of thirty children on each of the 231 traits of the California Behavior Inventory. The thirty children constituted the entire enrolment in the Institute's nursery school at that time. From one to two months after this task was completed, each teacher re-rated the children on each of 31 selected traits. Both the first ratings and the re-ratings were performed independently by each judge; and no recourse to original ratings was had at any time during the process of re-rating. All the ratings were made on a seven-point scale in accordance with rather detailed directions. For the reader's convenience, these directions are, except for certain minor omissions, reprinted verbatim herewith (2). It will be noted, from the directions, that the teachers are requested to rate each child *in comparison with others of the same age.* In addition, each nursery-school teacher is asked to indicate the confidence with which she made every rating for each child, and also to make a judgment as to whether, for the particular child being rated, the particular trait being rated was of "central or dominating importance" in the given child's personality.

A detailed description of the sample does not seem necessary for present purposes. It may suffice to remark that the clientele of the Institute's nursery school is drawn mainly from the upper-middle

and professional classes of Berkeley, California. According to such tests as the Minnesota Preschool Scale and the Merrill-Palmer, the average IQ of the nursery-school children falls usually between 115 and 120. Although no special clinical study is made of a child before admission to the nursery school, obvious or known "problem cases" are in general not accepted; on the whole, therefore, the nursery-school group may probably be considered as more or less "normal," with possibly a slight overweighting of children having mild behavior difficulties at home. Of the 30 children in the present study, 12 are boys, 18 are girls. As of March 1, 1930 (the time about which the ratings were made), the mean age of the boys was 41.4 months, the median, 42.5 months. The corresponding figures for the girls are 41.1 and 41.0. The boys ranged in age from 28 to 53 months (average deviation from the mean, 6.6 months); the girls, from 23 to 57 months (average deviation, 7.2 months).

B. DIRECTIONS FOR RATING

1. *The Traits and Symbols*

The Behavior Inventory consists of a list of numbered traits. Each trait in the Inventory has been described (*a*) by a suggestive heading; and (*b*) by an elaboration or illustration of the heading; in this elaboration, the *two extremes* of the trait are rather fully illustrated or defined, and the middle or *average* of the trait is usually stated somewhat more briefly.

Please give all children a rating for each trait, on a seven-point scale. Rate all the children *for each individual trait*, before proceeding to the next trait. Put your ratings on the Data Sheet which bears the *exact number* of the trait you are rating. Kindly notice that on each Data Sheet, the children have been arranged in order of chronological age. When rating, please take the child's chronological age into consideration. Each child should be rated only *in comparison with others of the same chronological age.*

The symbols to be used in rating are as follows:

"1": Child is *extreme and outstanding*, in the manner indicated in the *upper* third of the description of the trait. (May be thought of as +++)

"2": Child is *noticeably exceptional*, in the direction indicated by the *upper* third of the description of the trait. (May be thought of as ++)

"3": Child *differs from the average for his age,* leaning in the direction indicated by the *upper* third of the description of the trait. (May be thought of as +)

"4": *Average* for the child's age. (May be thought of as 0 or ±)

"5": Child *differs from the average for his age,* leaning in the direction indicated by the *lower* third of the description of the trait. (May be thought of as —)

"6": Child is *noticeably exceptional,* in the direction indicated by the *lower* third of the description of the trait. (May be thought of as — —)

"7": Child is *extreme and outstanding,* in the manner indicated in the *lower* third of the description of the trait. (May be thought of as — — —!)

A further way of amplifying the meaning of the symbols given above is as follows:

"1"—should be given to about 2 children in 100.
"2"—should be given to about 8 children in 100.
"3"—should be given to about 23 children in 100.
"4"—should be given to about 33 children in 100.
"5"—should be given to about 23 children in 100.
"6"—should be given to about 8 children in 100.
"7"—should be given to about 2 children in 100.

Notice that 80 per cent of the cases should fall within 3, 4, and 5. The remaining 20 per cent should be placed at the extremes—1, 2, 6, and 7. These figures apply to most (if not all) of the traits in the present rating scale.

Additional Symbols

a. Confidence of judgment. For each child, you are requested to record also the confidence with which you judge him in the particular trait being rated. If your rating in the particular trait was made with an amount of confidence *greater than average*, please underscore "3" opposite the child's name on the Data Sheet. If your rating was made with an *average amount of confidence*, underscore "2." If your rating was made with an amount of confidence *less than average*, please underscore "1."

b. Traits of central importance. If you consider certain traits

to be of central or dominating importance in a given child's personality make-up—that is to say, if you would give prominence to these traits in a description of the child's personality, or if you consider that these traits contribute in an outstanding way to the success or failure of the child's adjustment—then, when rating the child for these traits, *place a star or asterisk after your rating;* thus 6*, or 1*, etc.

c. Omitted ratings. If unable to rate a child, place a horizontal line in the box opposite the child's name (i.e., in the same box where the rating would ordinarily be put: thus, —).

2. *Cautions in Rating*

A warning may be offered concerning some possible sources of error in rating:

a. Rate each child with reference to others *of the same chronological age.* To illustrate, it would be obviously unfair to rate the self-reliance of a 21-month child by the same standard as a 48-month child.

b. Be sure to rate the child exclusively *on the trait under consideration.* Do not let the child's general personality, or your ratings on other traits, color your rating on any individual trait. Avoid *"halo-effect."*

c. State your honest opinion of a child on each specific trait—whether this opinion is complimentary or not. Too often, judges who are about to record an uncomplimentary rating, stop to think up possibilities of doubt against the adverse rating; and then they proceed to give the child "the benefit of the doubt." *This is not fair to the child about whom there is no doubt.* Please do not try to rationalize away a child's deficiencies, any more than you try to rationalize away its superiorities. The point is not to rationalize away either superiorities or inferiorities. Remember that your ratings of the child are not going to be used either for or against him; they are going to be used for scientific investigation exclusively. *Do not hesitate to put down uncomplimentary ratings, if these represent your opinion.*

d. Beware of the influence of the child previously rated upon your rating of the *present* child. If the preceding child was very bashful, for example, there is sometimes a tendency to rate the next child as quite bold. He may indeed be bold in comparison with the child just rated, but not in comparison with the average for his age.

The average for the child's age is, of course, the proper standard for comparison.

e. Some judges seem to be afraid to mark a child "average," apparently feeling that to place a child at "average" indicates some lack of discernment on the part of the rater. This is not true. Other judges, on the contrary, seem to feel that it is somehow "safest" to place a child at 4 (the average). This, also, is false. If the judge does not have the information necessary to rate a child on certain traits, then the only "safe" course is to refuse to rate the child on these particular traits.

C. THE SELECTION OF THE TRAITS TO BE RE-RATED

A word should be said with regard to the basis of selection of the 31 traits which were re-rated. Had unlimited time and patience been available, perhaps all 231 traits of the Behavior Inventory would have been re-rated. Under the practical conditions, however, such an undertaking was quite out of the question. The 31 traits finally selected were chosen partly because of their special interest to certain members of the Institute staff, and partly because of their representative character. In the Appendix, statistical data will be presented to show that, on the whole, the results for the 31 traits are, in many measurable respects, similar to results obtained from the full 231 traits of the Inventory. The body of the present monograph will be restricted to data obtained from the ratings and re-ratings of the 31 traits only.

II

THE CONSISTENCY OF TRAIT-RATINGS AND AGREEMENT BETWEEN JUDGES

A. Ratings of Thirty Children on Individual Traits

In this section we shall present verbatim each of the 31 selected traits.[1] Underneath each trait is listed certain statistical information for the trait; additional information is given in Tables 1 and 2. The meaning of the *notation* employed in the present section is as follows:

$r_{AA'}$ is the correlation between judge A's[2] ratings and re-ratings.

$r_{BB'}$ and $r_{CC'}$ are defined similarly to $r_{AA'}$.

$\bar{r}_{JJ'}$ is the mean of $r_{AA'}$, $r_{BB'}$, and $r_{CC'}$.[3] (The letter J in the subscript of $\bar{r}_{JJ'}$ stands for a single judge-in-general, unspecified as to identity.)

$r_{\overline{JJ'}}$ is the correlation between the *average* of judge A's, B's, and C's first ratings and the *average* of their re-ratings.

$\sigma_{AA'}$ is the mean of two standard deviations, viz., the SD of A's first ratings and the SD of her re-ratings.

$\sigma_{BB'}$ and $\sigma_{CC'}$ are defined similarly to $\sigma_{AA'}$.

---

[1]Occasionally in this presentation the heading of a trait has been slightly modified, in order to replace certain of the meaning lost by removing the trait from its context in the full Behavior Inventory (2).

[2]For convenience of symbolism, the three nursery-school teachers who performed the rating are designated as A, B, and C, respectively.

[3]The mean of the three coefficients provides a simple and direct measure of agreement. Theoretically it would be desirable, before averaging, to convert each coefficient to a value on a linear (i.e., equal-unit) scale. Such a procedure would, in general, tend to increase somewhat the values of the mean coefficient (cf. the discussion in Section C of the present chapter).

$\sigma_{88'}$ is the mean of two standard deviations, viz., the SD of the *average first rating* by A, B, and C and the SD of the *average re-rating* by A, B, and C.

$MD_{AA'}$ is the mean difference (without regard to sign) between A's first ratings and her re-ratings.

$MD_{BB'}$ and $MD_{CC'}$ are defined similarly to $MD_{AA'}$.

$MD_{88'}$ is the mean difference (without regard to sign) between the *average* of judge A's, B's, and C's first ratings and the *average* of their re-ratings.

In the remaining notation to be presented, only the teachers' *first* ratings are involved (as indicated by the absence of primes in the subscripts).

r_{AB} is the correlation between judge A's and judge B's ratings.

r_{AC} and r_{BC} are defined similarly to r_{AB}.

$\bar{r}_{J_1J_2}$ is the mean of r_{AB}, r_{AC}, and r_{BC}. (Cf. footnote 3, p. 12.)

$r_{8_18_2}$ is the (estimated) correlation between the *average* of judge A's B's, and C's ratings and the *average* of three *other* judges' ratings similar to A's, B's, and C's. (The formula used to compute $r_{8_18_2}$ is formula 154 of reference 5. In this formula, for our case, $a=b=3$; and we assume $\bar{r}_{pq}=\bar{r}_{PQ}=\bar{r}_{pQ}=1/3(r_{AB}+r_{AC}+r_{BC})$.

$r_{8\infty}$ is the estimated correlation between the average of judge A's, B's, and C's ratings and an infinite number of other judges' ratings similar to A's, B's, and C's. (The value of $r_{8\infty}$ is computed by the formula $r_{8\infty} = \dfrac{r_{8_18_2}}{\sqrt{r_{88'}}}$ —cf. formula 153*a*, reference 5). The symbols $r_{8_18_2}$ and $r_{88'}$ have already been defined above.

$r_{\infty\infty}$ is $r_{8_18_2}$ corrected for attenuation by the formula

$$r_{\infty\infty} = \frac{r_{8_18_2}}{\sqrt{r_{88'}}\ \sqrt{r_{88'}}} \text{ or } \frac{r_{8_18_2}}{r_{88'}}$$

σ_{AB} is the mean of two SD's, viz., the SD of A's ratings and the SD of B's ratings.

σ_{AC} and σ_{BC} are defined similarly to σ_{AB}.

MD_{AB} is the mean difference (without regard to sign) between A's and B's ratings of a trait.

MD_{AC} and MD_{BC} are defined similarly to MD_{AB}.

Each trait-description below is numbered as in the full California Behavior Inventory (2). The italic "*1*," "*4*," and "*7*" indicate the extreme, average, and opposite extreme of a trait, respectively. It may be noticed that, for some traits, the "average" behavior has been described comparatively briefly. The use of the California Behavior Inventory in general presupposes that the rater is sufficiently familiar with the behavior of children at the nursery-school ages so that an elaborate or detailed description of "average" behavior is as a rule unnecessary.

1. Amount of time spent in gross, overt activity—

1 Engaged in gross, overt activity a large portion of the time; child is characteristically active and in motion.

4 Engaged in gross, overt activity an average portion of the time.

7 Child is characteristically inactive or stationary; relatively seldom engaged in gross, overt activity.

$r_{JJ'}=.80$ $r_{SS'}=.94$ $r_{J_1J_2}=.70$ $r_{S_1S_2}=.91$

4.1 Inertia in getting started—

1 Very slow to get started overtly; slow to act; slow to express his desires or ideas in overt activity.

4 Average inertia in getting started.

7 Very quick to get started; quick to act; quick to express desires or ideas in overt activity.

$r_{JJ'}=.85$ $r_{SS'}=.94$ $r_{J_1J_2}=.70$ $r_{S_1S_2}=.91$

6. Vocalization—

1 Child is exceptionally active vocally; shouts, talks loud, sings songs, talks while playing or working, talks to others, etc. Much energy spent in vocal activity.

4 Average amount of vocal activity.

7 Child is exceptionally quiet; speaks in a low voice, rarely shouts, etc. Little energy spent in vocal activity.

$r_{JJ'}=.72$ $r_{SS'}=.85$ $r_{J_1J_2}=.60$ $r_{S_1S_2}=.89$

7.1 Overt emotionality: ease of stimulating to overt emotional response—

1 Child is *very easily* moved to overt emotional response. Highly

responsive. Mild unexpected stimuli, casual reproach, little accidents, and in general relatively weakly stimulating situations, readily produce definite overt emotion in the child.

4 Child is moved to overt expression of emotions with average difficulty or ease.

7 Child is *very difficult* to move to overt emotional response, being either indifferent, or (so far as overt indications are concerned) dully unresponsive. Exceptionally strong stimuli, and exceptionally pleasant (or unpleasant) situations are required to produce overt emotion in the child.

$\bar{r}_{JJ'}=.70$ $r_{33'}=.81$ $\bar{r}_{J_1J_2}=.49$ $r_{3\infty}=.82$

8.1 Inhibition of emotions: general statement—

1 Child customarily tries to control and restrain his emotions. Very highly inhibited, emotionally.

4 Average control or restraint of emotions. Average overtness of expression of emotions.

7 No restraint of emotions. Reaction to emotional stimuli is prompt, frank, free. Child rarely attempts to restrain, control, conceal, or inhibit his emotions.

$\bar{r}_{JJ'}=.70$ $r_{88'}=.78$ $\bar{r}_{J_1J_2}=.42$ $r_{8\infty}=.77$

9.2 General emotionality (implicit as well as overt): degree of emotional response—

1 Child's emotional response is customarily *very strong.* Child is intensely affected emotionally (whether he expresses his emotions openly and freely, or not).

4 Child's emotional responses are, in general, average in intensity.

7 Child's emotional response is characteristically small. Relatively strongly stimulating situations produce only mild or weak emotions in the child.

$\bar{r}_{JJ'}=.73$ $r_{88'}=.84$ $\bar{r}_{J_1J_2}=.62$ $r_{8\infty}=.91$

11. Recovery after emotional disturbance—

1 After the immediate emotional response the child behaves in its normal manner; no "after effects" of the emotion, such as unusual silence, inactivity, seclusion, showing off, pouting, brooding, irritability, etc. Child is exceptionally resilient, recovers very quickly.

4 Average resilience.

7 Lack of resilience. Child tends to remain (inwardly) in a persistently disturbed state; the immediate emotion brings an aftermath of disturbance in behavior (such as unusual silence, inactivity, seclusion, showing off, pouting, brooding, resentment, irritability, etc.). Internal emotion evidently continues, after

the emotional situation and the immediate emotional reaction have passed.

$r_{JJ'}=.64$ $r_{SS'}=.83$ $r_{J_1J_2}=.37$ $r_{Sx}=.70$

16. Predictability of emotions—

1 Emotions are very easily predicted; child's emotional response (or absence of emotional response) to a situation can be forecasted with little or no error.

4 Average. General nature of child's response can be predicted, more or less; but forecasts can be general only, or else must contain considerable error.

7 Emotions highly unpredictable; irregular, erratic emotionality; forecasts of emotional responses are very difficult to make.

$r_{JJ'}=.58$ $r_{SS'}=.74$ $r_{J_1J_2}=.33$ $r_{Sx}=.70$

21. Reaction to teasing—

1 Very sensitive, easily hurt, reacts strongly.

4 Average. Is sensitive to teasing, but behavior is not especially disorganized unless many children join in the teasing at one time.

7 Indifferent, quickly becomes callously disregardful.

$r_{JJ'}=.58$ $r_{SS'}=.79$ $r_{J_1J_2}=.20$ $r_{Sx}=.47$

24.1 Emotional reaction to sympathy or approval of adults—

1 Very sensitive; strong emotional response can be easily obtained by sympathy with child's efforts, or approval of accomplishment.

4 Average emotional responsiveness. Sympathy and approval are appreciated more when child is having difficulty, than at other times.

7 Emotionally insensitive or indifferent. Callous.

$r_{JJ'}=.36$ $r_{SS'}=.42$ $r_{J_1J_2}=.19$ $r_{Sx}=.63$

35.3 Mode of resistance to denial by another child: direct attempt to do or take what has been denied—

1 Child typically proceeds to try directly to do or take what he has been denied. E.g., if he has been denied a wagon (or a hammer, etc.), he proceeds at once to try to take it; if he wants (in a game) to play the father, he proceeds to do so, despite the previous objections of his playmate; etc.

4 Average attempt to override opposition. Attempts at direct fulfillment depend in part on the size of the other child, and the emphasis with which the other child has made the denial, etc.

7 Rarely attempts to ignore or override denial directly.

$r_{JJ'}=.80$ $r_{SS'}=.91$ $r_{J_1J_2}=.66$ $r_{Sx}=.90$

44.6 Detachment from other children —

1 Child disregards other children, pays little attention to others,

is often highly unaware of the presence of others.

4 Average. Notices others especially when seeking a playmate; disregards others in the midst of an earnest effort at painting, sawing, etc.; at other times gives ordinary notice and ordinary attention to others.

7 Child is keenly aware of the presence of others; detachment seems very difficult.

$\bar{r}_{JJ'}=.67$ $r_{BB'}=.81$ $\bar{r}_{J_1J_2}=.51$ $r_{B\infty}=.84$

46.32 Self-assertion: expression of his own desires or opinions in a group—

1 Quick to make plain his desires or opinions in a group. As quick and emphatic in making plain his displeasure and disapproval of what is being done or said, as to express approval and agreement. Markedly self-assertive.

4 Average. Less assertive in a group than with an individual playmate; but child indicates his desires and opinions to the group, when his feelings are strong.

7 Markedly inhibited or faltering in getting across his own desires and opinions. Typically submits, rather than making plain his own desires. Lacking in self-assertion.

$\bar{r}_{JJ'}=.76$ $r_{BB'}=.91$ $\bar{r}_{J_1J_2}=.59$ $r_{B\infty}=.85$

47.31 Retaliation: circuitous ("compensatory") retaliation: prompt—

1 Child often promptly "takes out" on another the offense done to himself; e.g., if hit by another, the child turns around and pushes or hits a third (innocent) child; if a toy has been taken away from him, the child turns around and takes a toy from a third (innocent) child, or hits a third child, or is otherwise nasty to him; etc.

4 Average. Child sometimes promptly "takes out" his grievance on an innocent third party, especially if deeply irritated, and if a smaller child is available to be "picked on."

7 Child never promptly "takes out" his grievance on any innocent parties, even if he is deeply affected and a smaller child is available as a victim.

$\bar{r}_{JJ'}=.83$ $r_{BB'}=.92$ $\bar{r}_{J_1J_2}=.70$ $r_{B\infty}=.92$

48.2 Initiative: need for outside suggestions and directions—

1 Child requires continuous outside suggestion and direction (from adults or other children) in order to choose an activity, or to select a method of overcoming an obstacle, etc. Finds it exceptionally difficult to get started and keep going "on his own hook." Lacks initiative.

4 Average initiative. Generally requires no suggestion or direc-

tion in order to get started; but if an unusual difficulty is encountered, the child is likely to be more or less at a loss, and require some outside suggestion or direction in order to do something about the obstacle or difficulty.

7 Child promptly and of own accord chooses the activity in which he is to engage, selects the method to be used in overcoming an obstacle, etc. Requires no outside suggestion or direction in order to get started and keep going. Has exceptional amount of initiative.

$r_{JJ'}=.75$ $r_{aa'}=.88$ $r_{J_1J_2}=.60$ $r_{a_1a_2}=.87$

51.1 Adjustability to new situations; emotional attitude toward new situations—

1 Child welcomes changes and new situations; is venturesome, exploring; enjoys novelty of a situation.

4 Average. Welcomes variations or small changes from the familiar, rather than genuinely new situations which he may not be prepared to meet.

7 Not at all venturesome; shrinks from making new adjustments; greatly prefers the habitual and familiar to the stress of reorganization required by the new. Routinated.

$r_{JJ'}=.76$ $r_{aa'}=.94$ $r_{J_1J_2}=.64$ $r_{a_1a_2}=.87$

56. Number of friends among children—

1 Child has many friends; "socially expansive."

4 Average number of friends.

7 Child has few friends; "socially constricted."

$r_{JJ'}=.79$ $r_{aa'}=.88$ $r_{J_1J_2}=.58$ $r_{a_1a_2}=.86$

65. Nervous habits —

1 Shows numerous nervous habits (or marked addiction to one nervous habit); e.g., fidgets, sucks thumb, bites nails, twitches face or shoulders, curls or twists hair, plays with mouth or face, etc.

4 Average. Shows some nervous habits under provocation.

7 Free from all signs of any nervous habits.

$r_{JJ'}=.51$ $r_{aa'}=.61$ $r_{J_1J_2}=.21$ $r_{a_1a_2}=.56$

68.2 Exhibitionism: through other than typical "showing off"—

1 Child (probably more or less deliberately) intrudes self into others' *attention*, by such devices as: explicitly calling attention to himself ("watch me!"); by boisterousness; by rapid running near others; by making a fuss over a slight injury; by causing a rumpus (e.g., through a quarrel); by mounting a high box, or standing on top of a jungle gym; by excessive affectionateness to others; by loquaciousness; etc.

4 Average. Appreciates attention, but seldom makes a special effort to attract it to himself.

7 Child is very unobtrusive and self-effacing; does not attract attention to himself.

$\bar{r}_{JJ'}=.68$ $r_{33'}=.88$ $\bar{r}_{J_1J_2}=.66$ $r_{3\infty}=.91$

87.1 Bashfulness with adults—

1 Self-conscious, shy, diffident, very easily embarrassed.

4 Average. Initial reaction to adult is one of some embarrassment, which, however, wears off fairly soon.

7 Rarely shows signs of embarrassment; not at all self-conscious; extremely self-composed; forward, "bold."

$\bar{r}_{JJ'}=.79$ $r_{33'}=.90$ $\bar{r}_{J_1J_2}=.53$ $r_{3\infty}=.81$

88. Apprehensiveness—

1 Anxious, frightened, or worried at the *prospect* of a situation which is new and unknown, unpleasant, or dangerous. E.g., child cries or is frightened at *prospect* of a physical examination, a vaccination, or a mental test; is scared at prospect of having to be away from familiar surroundings or from mother; is timid about making a new friend; etc.

4 Average. Becomes a little anxious, for example, while waiting for a vaccination, but does not usually worry much in advance. Occasionally slightly apprehensive at prospect of the unfamiliar.

7 Child never worries or becomes frightened in advance; is entirely carefree and unapprehensive.

$\bar{r}_{JJ'}=.84$ $r_{33'}=.96$ $\bar{r}_{J_1J_2}=.60$ $r_{3\infty}=.84$

91. Feelings of inferiority —

1 Child's average appearance and activity seem to indicate that he has rather deep-seated feelings of inferiority.

4 Average. Child seems to feel neither inferior nor superior.

7 Child's average appearance and activity seem to indicate that he has rather deep-seated feelings of superiority. Conceited.

$\bar{r}_{JJ'}=.50$ $r_{33'}=.83$ $\bar{r}_{J_1J_2}=.36$ $r_{3\infty}=.69$

100.2 Nervous laughter —

1 Child frequently makes use of a short, nervous laugh or giggle when nervous or embarrassed (especially in a social situation).

4 Average. Child occasionally reacts with a short giggle or laugh when embarrassed or nervous.

7 Child never makes use of a nervous laugh or giggle.

$\bar{r}_{JJ'}=.58$ $r_{33'}=.77$ $\bar{r}_{J_1J_2}=.18$ $r_{3\infty}=.44$

101. Number of interests —

1 Has many interests or favorite occupations.

4 Average. Has as many interests as most of the others of his age.

7 Has only a few interests; limited to a single favorite activity, or to inactivity.

$r_{JJ'}=.74$ $r_{BB'}=.87$ $r_{J_1J_2}=.58$ $r_{B_1B_2}=.86$

106. Onlooking—

1 Child is very often a spectator of a group, or of another child.

4 Average amount of spectator activity.

7 Child is rarely a spectator of a group, or of another child; either joins the child or group, or plays by self.

$r_{JJ'}=.87$ $r_{BB'}=.94$ $r_{J_1J_2}=.72$ $r_{B_1B_2}=.92$

107. Inertia of continued interest or occupation—

1 Child tends to continue automatically, by inertia, in the occupation (or lack of occupation) in which he finds himself. Lacks flexibility.

4 Average inertia; average shifting of occupation.

7 Child is extremely ready to alter his occupation or interest, upon suitable stimulation. Entirely free from inertia.

$r_{JJ'}=.78$ $r_{BB'}=.90$ $r_{J_1J_2}=.68$ $r_{B_1B_2}=.91$

112. Standards in work and play—

1 Child evidently has exceptionally high standards; tries to achieve a high standard of perfection in his work and play; tries to be thorough, and achieve a polish and finish. Likely to sacrifice speed for quality. Not at all sloppy, careless, rough-and-ready, uncritical, etc.

4 Average. Child's standards in work and play are not higher (nor lower) than those of others of his age. Child sacrifices thoroughness or polish when in a hurry.

7 Child has exceptionally low standards in his work and play; does not try to be thorough; neglects polish and finish in his activities; accepts low quality performance; is sloppy, or careless, or rough-and-ready, or uncritical, hasty, etc.

$r_{JJ'}=.62$ $r_{BB'}=.83$ $r_{J_1J_2}=.37$ $r_{B_1B_2}=.70$

117. Self-confidence—

1 Child feels cocksure of himself; is exceptionally self-confident even when exercising caution; over-confident, conceited.

4 Average. Self-confident in what he finds easy, becomes a little doubtful under difficulty.

7 Behavior is unsure, tentative, hesitant, faltering, inhibited. Child is markedly lacking in self-confidence.

$r_{JJ'}=.76$ $r_{BB'}=.87$ $r_{J_1J_2}=.59$ $r_{B_1B_2}=.86$

121. Ambition—

1 Child is ambitious in its undertakings. Is willing to try things that are hard to do; piles blocks exceptionally high, etc.

4 Average. Is fairly moderate in his undertakings and efforts.
7 Child is exceptionally unambitious. Characteristically undertakes only what is easy; is satisfied with a low pile of blocks, a half-filled wagon, a very moderate speed in pulling his express-wagon, etc.

$\bar{r}_{JJ'}=.82$ $r_{BB'}=.92$ $\bar{r}_{J_1J_2}=.69$ $r_{B\infty}=.91$

124. Imagination—
1 Child is *imaginative; e.g.,* impersonates his toys, plays many "make-believe" games, has imaginary playmates, or can even make up his own fairy tales and fantasies, etc.
4 Average. Plays "make-believe" games to some extent, but does not impersonate his toys much; has no imaginary playmates.
7 Child is unimaginative, matter-of-fact, "hard-boiled"; rarely plays "make-believe"; distinctly lacking in imagination.

$\bar{r}_{JJ'}=.66$ $r_{BB'}=.77$ $\bar{r}_{J_1J_2}=.69$ $r_{B\infty}=.992$

130. Use of facial expression: effectiveness—
1 Child's facial expressions are clear, readily understood, exceptionally effective in communicating feelings, wishes, invitations, etc.
4 Average. Facial expressions are clear and effective for major emotions, but less clear in communicating feelings, wishes, invitations, etc.
7 Child's facial expression is clear only in case of the gross, simple, major emotions. Facial expression (or lack of facial expression) is exceptionally ineffective or misleading in communicating shades of emotion, or wishes, invitations, etc.

$\bar{r}_{JJ'}=.71$ $r_{BB'}=.83$ $\bar{r}_{J_1J_2}=.49$ $r_{B\infty}=.81$

In Tables 1 and 2, we amplify or supplement the statistical data already presented for each trait in the immediately preceding pages. The traits in Tables 1 and 2 are numbered to correspond with the numbering of the traits just defined. The notation employed in the tables has already been fully explained on pages 12-13.

Perhaps the most impressive fact to be observed in Tables 1 and 2, and in pages 14-21, is the *wide difference in results for different traits*. While there is no guarantee that the same differences would be observed

TABLE 1

CONSISTENCY IN RATINGS OF EACH OF THIRTY-ONE TRAITS

Trait No.	Judge A $r_{AA'}$*	Judge A $MD_{AA'}$	Judge A $\sigma_{AA'}$	Judge B $r_{BB'}$	Judge B $MD_{BB'}$	Judge B $\sigma_{BB'}$	Judge C $r_{CC'}$	Judge C $MD_{CC'}$	Judge C $\sigma_{CC'}$	Composite judgment $r_{SS'}$	Composite judgment $MD_{SS'}$	Composite judgment $\sigma_{SS'}$
1	.79	.83	1.15	.81	.50	1.12	.81	.77	1.38	.94	.46	1.11
4.1	.85	.63	1.47	.94	.20	1.26	.76	.70	1.57	.94	.31	1.31
6	.82	.70	1.29	.66	.57	1.10	.68	.87	1.39	.85	.47	1.10
7.1	.48	1.00	1.30	.90	.23	1.09	.73	.87	1.34	.81	.48	.96
8.1	.80	.57	1.07	.81	.47	1.13	.50	1.13	1.54	.78	.54	1.03
9.2	.81	.57	1.19	.74	.47	1.08	.65	.93	1.40	.84	.50	1.08
11	.66	.73	.98	.83	.33	1.06	.44	.87	1.11	.83	.38	.82
16	.39	1.10	1.08	.85	.30	.87	.51	1.00	1.12	.74	.44	.75
21	.61	.70	1.03	.46	.67	.78	.66	.63	1.07	.79	.33	.72
24.1	.57	.93	1.03	.63	.63	.88	−.11	1.17	1.01	.42	.53	.70
35.3	.85	.67	1.30	.78	.53	1.24	.77	.83	1.56	.91	.52	1.22
44.6	.53	.73	1.07	.68	.67	1.20	.81	.80	1.46	.81	.49	1.05
46.32	.88	.80	1.39	.83	.53	1.13	.56	1.23	1.57	.91	.54	1.20
47.31	.85	.57	1.23	.78	.43	.99	.85	.42	.99	.92	.30	.93
48.2	.71	.80	1.18	.77	.47	1.00	.77	.67	1.25	.88	.38	1.01
51.1	.77	.70	1.38	.74	.63	1.22	.78	.77	1.48	.94	.34	1.22
56	.79	.63	1.24	.80	.53	1.17	.79	.37	1.01	.88	.44	[illegible]
65	.54	1.37	1.06	.66	1.03	1.06	.52	[illegible]	1.10	.61	[illegible]	[illegible]
68.2	.57	.77	1.07	.78	.57	1.18	[illegible]	.57	1.11	[illegible]	[illegible]	[illegible]
87.1	.82	.50	1.29	.81	.50	1.29	.73	[illegible]	1.29	[illegible]	[illegible]	[illegible]
88	.88	.67	1.42	.86	.41	1.15	.78	.63	[illegible]	[illegible]	[illegible]	[illegible]
91	.07	.67	.55	.80	.41	.96	.63	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]
100.2	.51	.53	.83	.73	.57	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]
101	.77	.63	1.15	.72	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]
106	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]
107	[illegible]	.65	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]
112	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]
[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]
[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]
[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]
[illegible]	[illegible]	.43	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]
Mean	[illegible]	[illegible]	1.12	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	1.24	[illegible]	.42	[illegible]

*For explanation of notation, see text. Sheppard's correction was not employed.

†If, instead of averaging the coefficients of correlation, we average the values of [illegible] corresponding [illegible] the mean value of [illegible] the figure [illegible]—which corresponds to a mean [illegible] of [illegible]

Trait No.	Judges A & B r_{AB}*	MD_{AB}	σ_{AB}	Judges A & C r_{AC}	MD_{AC}	σ_{AC}	Judges B & C r_{BC}	MD_{BC}	σ_{BC}	Composite judgment $r_{3_1 3_2}$	$r_{3\infty}$	$r_{\infty\infty}$
1	.74	.77	1.43	.65	1.10	1.59	.72	.87	1.42	.88	.91	.94
4.1	.80	.87	1.46	.60	1.20	1.65	.70	.87	1.46	.88	.91	.94
6	.75	.80	1.36	.52	1.03	1.54	.52	.97	1.27	.82	.89	.97
7.1	.58	1.03	1.29	.31	1.37	1.47	.58	.95	1.28	.74	.82	.91
8.1	.61	.87	1.21	.35	1.45	1.49	.30	1.30	1.40	.68	.77	.87
9.2	.70	.80	1.25	.74	.83	1.52	.42	1.23	1.33	.83	.91	.988
11	.51	.90	1.18	.18	1.17	1.23	.43	.87	1.15	.64	.70	.77
16	.33	1.03	1.12	.54	1.07	1.24	.12	1.17	.93	.60	.70	.81
21	—.12	1.07	.95	.49	.70	1.05	.22	.83	.85	.42	.47	.53
24.1	.24	.90	1.07	.16	1.17	1.17	.17	.80	.90	.41	.65	.98
35.3	.61	1.00	1.44	.76	1.03	1.69	.62	1.17	1.58	.86	.90	.95
44.6	.37	1.10	1.28	.64	1.13	1.52	.53	1.23	1.45	.76	.84	.94
46.52	.55	1.13	1.44	.63	1.30	1.75	.59	1.10	1.51	.81	.88	.89
47.51	.59	1.03	1.22	.73	.85	1.19	.78	.81	1.01	.88	.92	.96
48.2	.59	.93	1.14	.57	.87	1.24	.64	.87	1.13	.82	.87	.93
51.1	.60	.90	1.38	.65	1.07	1.56	.69	.83	1.35	.84	.87	.89
56	.54	1.00	1.32	.44	1.05	1.22	.77	.57	1.08	.81	.86	.92
65	.29	1.50	1.30	.13	1.80	1.35	.20	1.17	1.15	.44	.56	.72
68.2	.75	.60	1.15	.51	.73	1.06	.71	.60	1.08	.85	.91	.97
87.1	.52	.97	1.35	.58	1.10	1.39	.49	1.07	1.39	.77	.81	.86
88	.60	1.00	1.48	.67	1.07	1.49	.54	.95	1.27	.82	.84	.85
91	.53	.70	.96	.17	1.00	1.02	.38	1.05	1.13	.63	.69	.76
100.2	.26	.80	.97	—.07	1.43	1.09	.34	1.05	1.10	.39	.44	.51
101	.56	.87	1.18	.65	.83	1.30	.54	.90	1.17	.80	.86	.92
106	.74	.63	1.21	.76	.57	1.21	.66	.67	1.17	.89	.92	.95
107	.64	.93	1.20	.73	.70	1.42	.66	.90	1.28	.86	.91	.96
112	.34	.65	.82	.40	.63	.86	.37	.67	.79	.64	.70	.77
117	.64	.60	1.09	.43	.90	1.13	.69	.63	1.08	.81	.86	.93
121	.69	.67	1.20	.68	.80	1.26	.71	.73	1.25	.87	.91	.95
124	.74	.65	1.05	.67	1.00	1.02	.67	.57	1.02	.87	.992	1.15
130	.43	.83	1.05	.56	.63	.92	.48	.73	.88	.74	.81	.89
Mean	.54†	.89	1.21	.51†	1.02	1.31	.52†	.90	1.19	.74†	.81	.88

*For explanation of notation, see text. Sheppard's correction was not employed.

†If, instead of averaging the coefficients of correlation, we average the values of $\sqrt{1-r_{AB}^2}$ corresponding to r_{AB}, we obtain as the mean value of $\sqrt{1-r_{AB}^2}$ the figure .815—which corresponds to a mean r_{AB} of .58 (cf. reference 6). The mean values of r_{AC}, r_{BC}, $r_{3_1 3_2}$, $r_{3\infty}$, and $r_{\infty\infty}$, similarly computed, are .56, .57, .77, .84, and .91, respectively. See the discussion in Section C of this chapter. (In computing the mean value of $\sqrt{1-r^2}$, the value of the radical for the case in which $r_{\infty\infty}$ exceeds 1.00 was taken as 0; the value of the radical for the occasional case in which r is negative was taken as 1.00.)

in a larger sample of children, nevertheless it requires recognition that the striking differences observed between traits cannot be due, even in part, to ordinary errors of sampling—since our sample of children is *identical* for each trait. The results emphasize the desirability, in the present case, of thinking of reliability and validity in terms of specific traits rather than in terms of the total Behavior Inventory. The same observation is perhaps equally applicable to such composite traits as "introversion," "ascendance," etc. It is possible, moreover, that even the comparatively fine unit represented by a single Behavior Inventory trait requires further analysis; for discussion of this problem, however, the data of the present report are not adequate.

B. Ratings of Individual Children on Thirty-one Traits

It is desirable to consider individual differences among *children* in the same way that we have just considered individual differences among traits. Each child was rated on 31 traits, and then re-rated. The correlation between the ratings and re-ratings for a given child supplies a measure of the consistency with which that particular child was rated.

In Table 3, the consistency correlations of the nursery-school teachers' ratings are presented for each individual *child;* this table is analogous to Table 1, which showed the consistency of nursery-school teachers' ratings for each individual *trait*. The notation in Table 3 is to be read the same as that in Table 1.

TABLE 3
CONSISTENCY IN RATINGS OF EACH OF THIRTY CHILDREN

Child No.	Judge A $r_{AA'}$*	Judge A $MD_{AA'}$	Judge A $\sigma_{AA'}$	Judge B $r_{BB'}$	Judge B $MD_{BB'}$	Judge B $\sigma_{BB'}$	Judge C $r_{CC'}$	Judge C $MD_{CC'}$	Judge C $\sigma_{CC'}$	Composite judgment $r_{33'}$	Composite judgment $MD_{33'}$	Composite judgment $\sigma_{33'}$
1	.77	.29	.84	.56	.58	.85	.56	.93	1.39	.80	.42	.83
2	.52	.48	.67	.71	.45	.88	.48	.60	.88	.75	.38	.65
3	.79	.39	.87	.46	.45	.71	.49	1.00	1.25	.83	.34	.77
4	.74	.42	.88	.54	.45	.73	.51	.97	1.27	.84	.32	.79
5	.58	1.03	1.16	.70	.39	.78	.71	.68	1.11	.79	.46	.85
6	.80	.77	1.59	.90	.58	1.58	.86	.90	2.13	.94	.45	1.75
7	.10	.58	.59	.26	.39	.53	.42	1.13	1.10	.32	.31	.37
8	.78	.87	1.16	.89	.55	1.50	.75	.81	1.45	.92	.44	1.24
9	.75	.77	1.28	.52	.61	.91	.47	.90	1.18	.82	.44	.92
10	.80	.77	1.35	.74	.42	.92	.68	.84	1.58	.92	.33	1.05
11	.95	.48	1.51	.85	.48	1.25	.78	.61	1.54	.95	.33	1.23
12	.74	.42	.88	.79	.68	1.27	.69	.84	1.43	.83	.52	.98
13	.54	.74	.94	.57	.48	.79	.54	.74	.97	.74	.44	.72
14	.54	.94	1.15	.70	.58	1.04	.62	.90	1.41	.76	.63	1.02
15	.89	.87	2.07	.94	.52	2.24	.85	.71	2.15	.94	.48	2.04
16	.77	.77	1.04	.36	.52	.62	.53	.61	.82	.76	.40	.65
17	.32	.71	.83	.67	.42	.85	.36	.58	.75	.60	.46	.65
18	.51	.48	.69	.77	.39	.83	.71	.52	.89	.83	.27	.68
19	.69	.58	.91	.53	.55	.78	.60	.87	1.25	.76	.52	.81
20	.59	.58	.87	.40	.71	1.01	.49	.71	.90	.57	.43	.69
21	.73	1.23	1.58	.75	.68	1.24	.58	1.10	1.50	.81	.61	1.17
22	.75	.74	1.34	.79	.55	1.25	.67	.65	1.08	.84	.45	1.07
23	.49	.94	1.16	.59	.61	.94	.72	.65	1.10	.74	.41	.75
24	.54	.71	.96	.61	.61	.94	.38	.94	1.15	.77	.41	.76
25	.59	.32	.46	.23	.25	.38	.41	.55	.67	.42	.28	.32
26	.65	.77	1.05	.73	.45	.93	.67	.81	.99	.83	.48	.83
27	.49	.74	1.06	.71	.35	.76	.62	1.00	1.56	.70	.51	.91
28	.55	.77	.99	.78	.45	.83	.87	.45	1.25	.90	.41	.92
29	.72	.84	1.17	.60	.42	.77	.51	.55	.87	.74	.34	.64
30	.83	.71	1.84	.86	.32	1.07	.81	.68	1.42	.92	.42	1.26
Mean	.65†	.69	1.10	.65†	.50	.98	.61†	.77	1.21	.78†	.42	.91

*For explanation of notation, see text. Sheppard's correction was not employed.

†If, instead of averaging the coefficients of correlation, we average the values of $\sqrt{1-r_{AA'}}$ corresponding to $r_{AA'}$, we obtain as the mean value of $\sqrt{1-r_{AA'}}$ the figure .573—which corresponds to a mean $r_{AA'}$ of .67 (cf. reference 4, Table VIII). The mean values of $r_{BB'}$, $r_{CC'}$, and $r_{33'}$, similarly computed, are .67, .63, and .80, respectively. See the discussion in Section C of this chapter.

TABLE 4

Agreement between Judges in First Ratings of Each of Thirty Children

Child No.	Judges A & B r_{AB}*	MD_{AB}	σ_{AB}	Judges A & C r_{AC}	MD_{AC}	σ_{AC}	Judges B & C r_{BC}	MD_{BC}	σ_{BC}	Composite judgment r_{1}	r_{ss}	r_{oo}
1	.43	.77	.92	.28	1.27	1.24	.44	1.13	1.27	.65	.73	.81
2	.30	.68	.82	.62	.57	.88	.61	.60	.95	.76	.88	1.01
3	.23	.81	.86	.59	.97	1.24	.26	1.27	1.12	.63	.69	.76
4	.29	.81	.87	.56	.93	1.25	.32	1.17	1.12	.66	.72	.79
5	.34	1.16	1.11	.35	1.19	1.40	.63	.81	1.09	.70	.79	.89
6	.86	.61	1.74	.76	1.03	1.93	.81	1.06	1.86	.93	.96	.989
7	−.46	.87	.70	−.56	1.74	1.14	.15	1.32	1.05	..	..	..
8	.70	.97	1.61	.63	1.10	1.52	.73	.97	1.55	.87	.91	.95
9	.49	1.00	1.20	.31	1.29	1.33	.33	.87	1.05	.64	.71	.78
10	.57	1.10	1.26	.78	.84	1.47	.61	.90	1.12	.85	.89	.92
11	.82	.81	1.49	.67	1.03	1.62	.53	1.00	1.34	.86	.88	.91
12	.64	.87	1.23	.58	.97	1.31	.41	1.26	1.51	.78	.86	.94
13	.51	.84	1.02	.37	.94	1.04	.57	.48	.91	.73	.85	.987
14	.71	.84	1.22	.64	.84	1.46	.65	.77	1.24	.88	.987	1.11
15	.84	.77	2.31	.87	.87	2.28	[illegible]	1.15	2.22	.93	.96	.989
16	.47	1.00	1.03	.49	.97	1.04	.13	.68	.71	.63	.72	.81
17	.58	.61	.90	.44	.68	.81	.34	.71	.77	[illegible]	.92	[illegible]
18	.50	.71	.80	.54	.48	.78	.70	.42	.78	[illegible]	.89	.98
19	.61	.61	.97	[illegible]	[illegible]	1.17	.72	.65	1.35	.92	.94	1.05
20	.17	1.00	.98	.14	.87	.93	.55	.65	.92	[illegible]	[illegible]	[illegible]
21	.46	1.52	1.73	.43	[illegible]	[illegible]	.58	1.39	[illegible]	[illegible]	[illegible]	[illegible]
22	.67	.81	1.44	.58	[illegible]	[illegible]	.57	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]
23	−.12	1.29	1.14	.26	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]
24	.37	.54	1.11	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]
25	.32	.52	.53	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]
26	.55	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]
27	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]
28	[illegible]	.71	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]
29	[illegible]	[illegible]	[illegible]	−[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]
30	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	1.34	[illegible]	[illegible]	[illegible]
Mean	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	.91	[illegible]	[illegible]	[illegible]	[illegible]

* For explanation of notation, see text. Sheppard's correction was not employed.

† If, instead of averaging the coefficients of correlation, we average the values of $\sqrt{1-r_{AB}^2}$ corresponding to r_{AB}, we obtain as the mean value of $\sqrt{1-r_{AB}^2}$ the figure [illegible]—which corresponds to a mean r_{AB} of [illegible] [illegible]

thus $r_{AA'}$ signifies the correlation between judge A's first ratings and her re-ratings; $MD_{AA'}$ signifies the mean difference (without regard to sign) between A's first ratings and her re-ratings; etc. (cf. pp. 12-13). Inspection of Table 3 shows that the differences in results for different children are fully as great as those for different traits.

Turning now to a consideration of the agreement between various teachers (as contrasted with the consistency of a given teacher), we find that here, too, the agreement varies greatly among the children. In Table 4 we find, for example, that judge A's ratings of child No. 11 on 31 traits correlate with judge B's ratings of the same child on the same traits to the extent of $r_{AB} = .82$; for child No. 10, however, the agreement is much lower, viz., $r_{AB} = .57$. Table 4, giving the statistical information for individual children, is analogous to Table 2, giving the corresponding information for individual traits. Inspection of Table 4 shows that the differences in results for different children are fully as great as those for different traits.

The wide individual differences among children observed in Tables 3 and 4 are in our opinion among the most significant findings of the present investigation. We have seen many studies which emphasize the differing capacities of judges to rate and the differing susceptibility of traits to rating. The equally important fact that there are great individual differences in the accuracy with which different persons can be judged has not (outside of clinical circles) received the attention which it requires. If we may accept the

results in Tables 3 and 4, it appears that the reliability and validity of rating are as much a function of the sample of *subjects* as of raters, traits, conditions of rating, etc. We need intensive investigation of the factors responsible for individual differences in "ratability."

C. The Age Factor

A comparison of the mean figures at the bottom of Tables 1 and 2 with the mean figures at the bottom of Tables 3 and 4 is of special interest. It might be suspected that the coefficients of correlation in Tables 1 and 2 are affected by a spurious age factor (in spite of the direction to the teachers to rate each child in comparison with others of the same age—cf. page 7). If such a spurious factor were operating, its effect would be to inflate the r's of Tables 1 and 2. In Tables 3 and 4, however, the r's are surely not so affected, since every r in these tables is based on ratings of a *single child* only; so far as each r in these tables is concerned, age has been "held constant." Hence, we should expect, in the event of a spurious age factor, that the r's in Tables 3 and 4 would be somewhat lower than those in Tables 1 and 2. And this is actually the case; but the difference in general is not great; and there are other possible explanations for the difference, besides the age factor.

One of the other possible explanations is based on the fact that high coefficients of correlation tend to be associated with high standard deviations (5, p. 225). Thus, it is conceivable that the ratings for one child may hover around 3, with a comparatively low standard deviation; the ratings for another may hover around 4, with

a low standard deviation; etc. When, however, the ratings for *all* children on a single trait are involved, the standard deviation of the ratings may be comparatively high. Examination of the average σ's at the bottom of Tables 1-4 suggests that this factor, while present, is influential to only a slight degree.

Another possible explanation of the difference in mean r's of Tables 1 and 3 and of Tables 2 and 4 is that, while regression is linear in the correlations of Tables 1 and 2, regression departs from linearity in Tables 3 and 4. Inspection of the correlation charts from which the r's were computed fails, however, to substantiate this possibility.

The other possible explanation of the difference in mean r's of Tables 1 and 3 and of Tables 2 and 4 rests on the fact that a difference between r's of (say) .70 and .80 is not, in any important sense, equal to a difference between r's of (say) .10 and .20. Suppose, for example, that the r's for individual *children* vary more than the r's for individual *traits*. Then the *mean* correlation for all the individual children may be spuriously low in comparison with the *mean* correlation for all the individual traits. It is desirable, therefore, to convert r to a linear scale. When re-measurements are involved, this is customarily done by substituting $\sqrt{1-r}$ for r.[4] We have made this substitution for the r's in Tables 1 and 3, with results as stated in footnotes to these tables. The figures for the theoretically proper mean[5] values of r are in general slightly higher than the uncorrected mean values; but the *differences* between the corrected mean values in Tables 1 and 3 are only slightly smaller than those between the uncorrected.

The same type of analysis was continued for the r's of Tables 2 and 4, except that here we have used $\sqrt{1-r^2}$ [6] rather than $\sqrt{1-r}$ to convert r to a linear scale. The use of $\sqrt{1-r^2}$ is customary when comparison is made not between re-measurements,

[4]This substitution is based on the fact that the standard error of measurement equals (in Garrett's notation) $\sigma_1\sqrt{1-r_{12}}$. Cf. reference 3, p. 275.

[5]The unweighted mean was used because n, the number of cases, is constant, and because the refinement of weighting each r inversely as the square of its probable error was considered probably superfluous.

[6]The coefficient of alienation (5, p. 174).

but between measurements by different instruments (such as different intelligence tests, or different judges, etc.). The theoretically proper mean values again exceed the uncorrected means, this time quite definitely. Moreover, the differences between the theoretically proper mean values for Tables 2 and 4 are definitely smaller than the differences between the uncorrected mean values given in these tables. This suggests that, so far as the agreement between different judges' ratings of individual traits is concerned, the age factor is less important than an uncritical comparison of Tables 2 and 4 might suggest.

It must be remembered that, at the very most, the age factor can explain only the *difference* between the mean r's at the bottom of Tables 1 and 3 and of Tables 2 and 4. The age factor cannot, in general, account for any but a small fraction of the degree of relationship shown for individual traits in Tables 1 and 2.

SUMMARY

The present section is concerned with the consistency of nursery-school teachers' ratings and the agreement between teachers. The unit of statistical treatment is the individual trait, and the individual child. Correlation coefficients for the individual traits were obtained by correlating the ratings for the various *children* on the particular *trait* in question. Correlation coefficients for the individual children were obtained by correlating the ratings on the various *traits* for the particular *child* in question.

For individual traits, the consistency correlation between a single judge's ratings and re-ratings, made from one to two months apart, is about .70. There are, however, wide differences in results for different traits. If the ratings of three teachers are pooled, the

consistency correlation comes to about .85 on the average, again with wide individual differences among the various traits. The detailed presentation of these facts is given in Table 1.

The agreement, in terms of correlation, between one judge's ratings of a trait and another's averages somewhat over .55; the agreement between the pooled ratings of three teachers and an infinite number of other teachers is estimated, on the average, as over .80. Certain traits are rated with far more agreement than others (cf. Table 2).

For individual children, the consistency correlation between ratings and re-ratings, made from one to two months apart by a single judge, averages about .65; this rises to .80 if ratings by three teachers are pooled (cf. Table 3). The correlation between one judge's ratings of a child and another's averages over .50; the agreement between the pooled ratings of three teachers and an infinite number of other teachers is estimated as over .80 (cf. Table 4).

The differences in results for different children are fully as great as the differences in results for different traits. It is therefore evident that the reliability and validity of rating are as much a function of the sample of *subjects* as of traits, raters, conditions of rating, etc. This is a point which does not, as yet, seem to have received adequate emphasis.

A study of the possible influence of a spurious age factor in the correlations for the individual traits indicated that in the present instance the effect of this *factor must, in general, be small.*

The discussion in the present summary and in the text has been mainly in terms of the coefficient of correlation. For those who prefer it, however, the average difference between judges' ratings has also been included in all the tables of the text.

III

CONSISTENCY IN THE "STARRING" OF TRAIT-RATINGS

Perhaps the most interesting feature of the present study is the "starring" of ratings for outstanding traits. It will be recalled that each nursery-school teacher was requested to affix an asterisk or star to her rating of every trait which, for the particular child being rated, she considered of "central or dominating importance" in that child's personality. Is a nursery-school teacher able to make, with reasonable consistency, this rather fine judgment of the importance of an individual trait for an individual child? The answer to this question, for each teacher separately, is contained in Table 5.

Table 5 is clear enough in its general implication. For purposes of discussion, we shall restrict ourselves to the data in Table 5 for judge B. In this particular table it is evident that the actual frequencies in each cell of the four-fold table depart definitely from what would be expected by the operation of pure chance. Nevertheless, inspection merely of the marginal frequencies is sufficient to show that judge B's consistency is far from perfect. With perfect consistency, the marginal frequencies for B's re-ratings would be 127 and 803, instead of (as they are) 96 and 834. Judge B starred 31 more of her first ratings than of her re-ratings; it is evident that the standard for inclusion in the starred category was more lenient for the first ratings than for the second. Accepting this change of standard, we may now properly inquire, what is the highest coefficient of contingency which it is possible to find under the condition of this particular change of standard? The answer is given in Table 5A, which is an artificial or hypothetical table with the same *marginal* frequencies as in Table 5 for judge B, but with *cell*-frequencies adjusted to yield the highest possible contingency coefficient. The coefficient of mean square contingency for

TABLE 5

CONSISTENCY OF STARRING OF TRAIT-RATINGS

Re-ratings, A	First Ratings: Judge A Starred	Unstarred	Total
Unstarred	162 (210)*	680 (632)	842
Starred	70 (22)	18 (66)	88
Total	232	698	930

Coef. of mean square contingency.... .38†
Coef. of tetrachoric correlation.... .73

Re-ratings, B	First Ratings: Judge B Starred	Unstarred	Total
Unstarred	65 (114)*	769 (720)	834
Starred	62 (13)	34 (83)	96
Total	127	803	930

Coef. of mean square contingency.... .45†
Coef. of tetrachoric correlation.... .79

Re-ratings, C	First Ratings: Judge C Starred	Unstarred	Total
Unstarred	75 (113)*	704 (666)	779
Starred	59 (21)	88 (126)	147
Total	134	792	926

Coef. of mean square contingency.... .30†
Coef. of tetrachoric correlation.... .56

*The figures in parentheses indicate the cell-frequency to be expected by chance.

†The highest possible coefficient of contingency in a 2 x 2 table is .707. Due to asymmetry of the marginal distributions, the highest possible C for judge A is .49; for B, .65; for C, .69.

TABLE 5A

HIGHEST POSSIBLE COEFFICIENT OF CONTINGENCY FOR JUDGE B (CF. TABLE 5)

First Ratings

Re-ratings	Starred	Unstarred	Total
Unstarred	31	803	834
Starred	96	0	96
Total	127	803	930

Coef. of mean-square contingency, for the hypothetical Table 5 A .... .65

Table 5A is .65. The highest possible coefficient of contingency, in a symmetrical four-fold table, is .707 (7, p. 66). The difference, then, between .707 and .65, is a measure of the effect of judge B's inconsistency of standard for starring. But the coefficient of contingency for judge B in Table 5 is not .65, it is .45. The difference between .65 and .45 is a measure of judge B's inconsistency, apart from the inconsistency of her standard for inclusion of ratings in the starred category.

In addition to the difference between the highest possible and the actual contingency coefficient, use has been made of the coefficient of tetrachoric correlation (1). This coefficient, for judge B, equals .79.

The data for judges A and C have been given the same statistical treatment as outlined above for judge B. The results are recorded in Table 5.

Table 5 does not, on the whole, encourage very much confidence in a single judge's evaluation of the importance of a particular trait in the personality of a

particular child.[7] But is it still possible that the

TABLE 6

CONSISTENCY OF "SIGNIFICANCE SCORES" FOR RATINGS OF INDIVIDUAL CHILDREN ON INDIVIDUAL TRAITS

Re-ratings: No. of judges starring their ratings of a trait for a child	First Ratings: No. of judges starring their ratings of a trait for a child: 0	1	2	3	Total
3	1 (25)*	7 (9)	16 (4)	15 (1)	39
2	5 (25)	11 (10)	15 (4)	9 (1)	40
1	50 (85)	48 (32)	31 (12)	5 (5)	134
0	531 (453)	159 (173)	24 (66)	3 (15)	717
Total	587	225	86	32	930

Coef. of mean square contingency........ .55
r (computed from the 4 x 4 table by the usual product-moment formula)........ .60

*The figures in parentheses indicate the cell-frequency to be expected by chance.

[7]Conceivably our unit—the individual trait-rating—is too fine to reveal the full extent of the teachers' consistency. A teacher might, for example, star her rating for "physical activity" at one time; and at another time, star her rating for "number of interests." The result would constitute a statistical inconsistency, although actually the inconsistency might be comparatively supercificial. The discrepancy between the coefficients of correlation in Table 7 and in Table 8 may perhaps be regarded as supporting this point of view. (Tables 11 and 12, in the succeeding section, show similarly discrepant coefficients.)

consensus of judgment by three teachers is sufficiently consistent to deserve more serious attention. Suppose, for example, that two of the three teachers starred their ratings of a certain trait for a certain child. So far as the importance of this trait for the particular child is concerned, the "significance score" may be considered as 2. Upon re-rating, will the significance score also equal 2? The general answer to this correlational question is given in the 4 x 4 table on page 36 (Table 6). In Table 6, the number of times a given child's ratings on a given trait were starred at the first rating is given on the X-axis; the number of times a given child's ratings on a given trait were starred at the re-rating is given on the Y-axis. The coefficient of mean square contingency, for Table 6, is .55; the highest possible contingency coefficient for a 4 x 4 table in which consistency is perfect would be .866[8] (7, p. 66). As a matter of interest, the coefficient of correlation for Table 6, computed from the 4 x 4 chart by the usual product-moment formula, is .60.

Our unit of discussion, up to this point, has been the *trait-child*—i.e., we have studied the consistency of judges' starring of ratings of *individual* traits for *individual* children. We plan, now, to broaden this unit. The question to be answered is: How many times were the ratings of (say) trait No. 1 starred for the *entire group* of children? What is the consistency of

[8]The highest possible contingency coefficient for a 4 x 4 table with exactly the same marginal frequencies as those of Table 6 is .79.

this number from first rating to re-rating? Table 7 presents the facts. For any individual teacher, the consistency is comparatively low—the highest r is .56.

TABLE 7

CONSISTENCY WITH REGARD TO THE NUMBER OF TIMES THE RATINGS OF EACH TRAIT WERE STARRED

Judge	Variable	Mean	Median	S.D.	r
A	No. of starred first ratings	7.5	7.3	4.0	.56
	No. of starred re-ratings	2.8	2.4	2.3	
B	No. of starred first ratings	4.1	4.1	2.5	.47
	No. of starred re-ratings	3.1	3.0	2.5	
C	No. of starred first ratings	4.3	2.3	4.5	.49
	No. of starred re-ratings	4.7	4.4	3.5	
A, B, & C, combined*	No. of starred first ratings	5.3	4.7	2.9	.71
	No. of starred re-ratings	3.6	3.7	3.1	

*Mean number of starred ratings for each trait by A, B, and C.

For the combined judgment of the three teachers, however, the coefficient of correlation is .71. The consistency correlations would, of course, be expected to be somewhat higher if the number of children being rated were larger.

Let us shift the unit of discussion, now, from *trait* to *child*. The ratings of some children are starred far more often than the ratings of others. For example, in the case of the re-ratings for a certain child, only *one* of the three teachers starred *one* of her trait-ratings for this child; whereas, for another child, *no* teacher starred fewer than 16 of her 31 re-ratings for the child. Are the nursery-school teachers consistent in the number of times they star their trait-ratings for a child? The coefficients of correlation in Table 8 give the facts. It will be noticed that the r's in Table 8 are definitely higher than the r's in Table 7. Part of this, probably,

is due to the somewhat greater individual differences (higher S.D.'s) among children than among traits (*cf.* Tables 7 and 8).

TABLE 8

CONSISTENCY WITH REGARD TO THE NUMBER OF TIMES THE TRAIT-RATINGS OF EACH CHILD WERE STARRED

Judge	Variable	Mean	Median	S.D.	*r*
A	No. of 1st ratings starred	7.7	7.5	5.6	.76
	No. of re-ratings starred	2.9	1.7	4.1	
B	No. of 1st ratings starred	4.2	2.1	5.1	.86
	No. of re-ratings starred	3.2	1.5	4.8	
C	No. of 1st ratings starred	4.5	4.7	3.2	.74
	No. of re-ratings starred	4.9	3.3	4.5	
A, B, & C, combined*	No. of 1st ratings starred	5.5	4.5	2.0	.89
	No. of re-ratings starred	3.7	2.5	2.0	

*Mean number of traits starred for each child by A, B, and C. (The statistical constants given in the last two lines of the table were computed from grouped data; the others, from ungrouped data.)

A summary and evaluation of the data in Tables 5-8 are not easy. Persons with different standards of statistical judgment will come to different conclusions. To some, the coefficients for consistency of starring will appear rather low, or even very low. Others, on the contrary, may feel reassured from the fact that the degree of consistency found is certainly well above that due to chance (in one instance, as high as $r = .89$); and also from the observation that conventional statistical treatment may not reveal the full extent of the consistency that really exists. For ourselves, we regard the obtained coefficients of consistency as indicating, at the very least, good research possibilities.

The study of the starring of trait-ratings has been restricted to an investigation of self-agreement, or consistency. The problem of inter-judge agreement

in the starring of trait-ratings will be reserved until results for the full 231 traits of the Behavior Inventory become available.

SUMMARY

The directions for the California Behavior Inventory require the nursery-school teacher to affix an asterisk or star to the rating of any trait which, for the particular child being rated, she considers of special significance in that child's personality. The present section is devoted to a study of the consistency of evaluating trait-importance, or in other words the consistency of "starring." For this purpose, three nursery-school teachers' ratings and re-ratings of 30 children on 31 traits are available.

The consistency of a single teacher's starring of individual trait-ratings is too low to encourage confidence. The pooled judgment of the three teachers is more satisfactory, but still far below conventional standards for a complete mental test. As measured by the coefficient of contingency from a 4 x 4 table, the consistency of the pooled judgment is .55. (The highest possible coefficient of contingency in a symmetrical 4 x 4 table is .866.)

To evaluate the importance of an individual trait for an individual child requires a rather minute judgment; to evaluate the general importance of an individual trait for an entire group of 30 children should be easier. Each teacher's evaluation of the general importance of a trait was obtained simply by counting the total number of times she starred her ratings of the

trait. This number was correlated with the number of times she starred her re-ratings of the trait. The correlation, for the individual teacher, averages no higher than around .50. For the pooled judgment of three teachers, the consistency correlation is slightly over .70. The consistency correlations would, of course, be somewhat higher if the number of children rated were greater than 30.

A companion problem to the one just discussed is the consistency of the total number of trait-ratings starred for each individual *child*. The consistency correlations for this, interestingly enough, are much higher than for the individual traits—being around .79 for an individual teacher, and .89 when the judgments of the three teachers are pooled.

IV
THE CONSISTENCY OF CONFIDENCE-RATINGS

The nursery-school teachers were asked to indicate on a three-point scale the degree of confidence with which they rated each trait for each child.[9] For convenience in discussion, the confidence-ratings accompanying the first trait-ratings will be termed "initial confidence-ratings," and the confidence-ratings accompanying the re-ratings will be termed "confidence–re-ratings." The same questions present themselves concerning the confidence-ratings as have been answered in the previous section on the starring of trait-ratings; and a similar technique of investigation and presentation will here be employed.

The upper portion of Table 9 shows the agreement between judge A's initial confidence-ratings and her confidence–re-ratings. The total number of confidence-ratings by A, for which confidence–re-ratings were made, is 846.[10] Using the coefficient of mean square contingency as the measure of relationship, we find that judge A's consistency equals a coefficient of .42. The highest possible coefficient of contingency in a symmetrical 3 x 3 table is .816 (7, p. 66); the highest

[9]One of the nursery-school teachers (judge C) made her confidence-ratings primarily on the basis of the trait being rated, without taking into account *also the child being* rated. Since it was desired that the work of rating should, so far as possible, be voluntary and self-motivated, *no objection* to this method was interposed.

[10]Differing from 930, due to occasional omissions of confidence-ratings.

TABLE 9
CONSISTENCY OF CONFIDENCE-RATINGS

Confidence-Rating: Judge A

Confidence-Rating - A	Low	Average	High	Total
High	61 (96)*	258 (262)	206 (168)	525
Average	52 (50)	163 (138)	61 (88)	276
Low	41 (8)	1 (22)	3 (14)	45
Total	154	422	270	846

Coefficient of mean square contingency.......... .42†

Confidence-Rating: Judge B

Confidence-Rating - B	Low	Average	High	Total
High	11 (24)*	83 (114)	76 (34)	170
Average	77 (83)	454 (417)	88 (119)	619
Low	27 (8)	33 (41)	1 (12)	61
Total	115	570	165	850

Coefficient of mean square contingency.......... .57†

Confidence-Rating: Judge C

Confidence-Rating - C	Low	Average	High	Total
High	1 (2)*	5 (3)	0 (1)	6
Average	260 (254)	271 (281)	74 (71)	605
Low	93 (96)	114 (105)	25 (27)	232
Total	354	390	99	843

Coefficient of mean square contingency.......... .50†

*The figures in parentheses indicate the cell-frequency to be expected by chance.

†The highest possible coefficient of contingency in a 3 x 3 table is .816. Due to asymmetry of the marginal distributions, the highest possible C for judge A is .62; for judge B, .76; for judge C, .60.

possible coefficient for a 3 x 3 table with the marginal frequencies in Table 9 is .62 (cf. Table 5, *circa*).

The data for the consistency of confidence-ratings by judge B and judge C are also presented in Table 9, similarly to that for judge A. The coefficient of mean square contingency for judge B is .37; for judge C, .09. Evidently, as in the case of the starring of trait-ratings, the consistency of the confidence-ratings for individual trait-ratings of individual children is low. This is particularly true of the confidence-ratings by judge C.

Omitting judge C's confidence-ratings, we have investigated the consistency of judge A's and B's combined confidence-ratings. Suppose, for example, that both judge A and judge B rated a certain child on trait No. 1 with high confidence, and re-rated this child on this trait with the same degree of confidence. This agreement would be tallied in the "High-High" cell of Table 10. If judge A rated a child on a trait with average confidence, while judge B rated this child with high confidence, the combined confidence-rating is "above average"; if the judges re-rated this child on the trait with the same degree of confidence as in the first rating, then this agreement would be tallied in the "Above Av.–Above-Av." cell of Table 10. The coefficient of contingency for Table 10 is .49. The highest possible coefficient of contingency in a symmetrical 5 x 5 table is .894; the highest possible coefficient in a table with the marginal frequencies of Table 10 is .82. As a matter of interest, the product-moment correlation for Table 10, computed

TABLE 10

CONSISTENCY OF AVERAGE CONFIDENCE-RATINGS OF JUDGES A AND B

Average Confidence-Rating

Average Confidence Re-rating	Low	Below Av.	Av.	Above Av.	High	Total
High	0 (8)*	11 (29)	34 (71)	57 (47)	80 (27)	182
Above Av.	14 (19)	69 (71)	188 (174)	134 (116)	42 (67)	447
Av.	15 (11)	38 (39)	124 (95)	51 (64)	17 (37)	245
Below Av.	7 (2)	14 (6)	15 (14)	0 (9)	0 (5)	36
Low	4 (1)	15 (3)	1 (8)	0 (5)	0 (3)	20
Total	40	147	362	242	139	930

Coef. of mean square contingency........ .49
r (computed from the 5 x 5 table by the usual product-moment formula)......... .44

*The figures in parentheses indicate the cell-frequency to be expected by chance.

from the 5 x 5 chart by the usual formula, is .44.

Different traits and different children are rated with varying degrees of confidence by the judges. For each trait the unweighted average confidence with which it was rated for all 30 children was correlated with the unweighted average confidence at the re-rat-

ing.[11] The results, shown in Table 11, may well be considered disappointing. Similar correlations, how-

TABLE 11

CONSISTENCY OF TEACHERS' AVERAGE CONFIDENCE-RATINGS FOR EACH OF THIRTY-ONE TRAITS

Judge	Variable	Mean*	Median*	S.D.*	r
A	Conf-rating	2.22	2.10	.38	.36
	Conf.-re-rating	2.39	2.38	.42	
B	Conf.-rating	2.12	2.13	.20	.45
	Conf.-re-rating	2.19	2.21	.20	
A & B, combined†	Conf.-rating	2.20	2.21	.23	.44
	Conf.-re-rating	2.42	2.42	.26	

Because of judge C's method of making confidence-ratings (cf. p. 42), her average confidence-ratings for different traits were, in general, either 1, 2, or 3, without any graduation. The coefficient of contingency for judge C's consistency of confidence in the rating of each of the 31 traits is .19.

*Counting low confidence as 1, average confidence as 2, and high confidence as 3 (footnote 11).

†Mean of A's and B's average confidence-ratings for each trait. (Cf. text.)

ever, between the average confidence with which each *child* was rated and re-rated on all 31 traits by each judge[12] are comparatively favorable (averaging around .80—cf. Table 12).

The results of the present section concern only the consistency or self-agreement of confidence-ratings, giving no attention to the question of inter-judge agreement. The problem of inter-judge agreement in confidence-ratings, like that of inter-judge agreement in the starring of trait-ratings, is reserved for study when results for the full 231 traits of the Behavior Inventory become available.

[11]*In* obtaining the average confidence-rating for each trait, "high confidence" was counted as 3, "average" as 2, and "low confidence" as 1. It is, of course, possible to calibrate each teacher's confidence-ratings on the basis of the *normal* curve. But the correlation between average-confidence-ratings based on raw values, and average-confidence-ratings based on calibrated values, turns out to be very high (over .990), thus rendering the use of calibrated values unnecessary.

[12]Since judge C's confidence-ratings were, in general, made without regard to the individual child being rated (p. 42), we have excluded her confidence-ratings from the present treatment.

TABLE 12

CONSISTENCY OF TEACHERS' AVERAGE CONFIDENCE-RATINGS FOR EACH OF THIRTY CHILDREN

Judge	Variable	Mean*	Median*	S.D.*	r
A	Conf.-rating	2.19	2.20	.28	.74
	Conf.-re-rating	2.60	2.65	.18	
B	Conf.-rating	2.12	2.16	.23	.84
	Conf.-re-rating	2.19	2.18	.23	
A & B, combined†	Conf.-rating	2.18	2.21	.24	.80
	Conf.-re-rating	2.39	2.41	.18	

*Counting low confidence as 1, average confidence as 2, and high confidence as 3 (p. 46).

†Mean of A's and B's average confidence-ratings for each trait. (Cf. text.)

SUMMARY

The consistency of confidence-ratings has been studied by methods similar to those already used in connection with the starring of trait-ratings. For the average confidence with which two judges rate and re-rate all the children on all the traits (n=930), the coefficient of contingency is about .50. The consistency of the average confidence with which two teachers rate a single trait for all children is expressed by a coefficient of correlation of about .45 (n=31 traits); this is surprisingly low, in view of the contingency coefficient cited above. If, however, the confidence-ratings for a given child on all the traits are averaged, the consistency found is moderately high (r=.74 for judge A, .84 for judge B; n=30 children). It appears, then, that in general the judges are moderately consistent in the average confidence of their ratings for children; but the distribution of their confidence among the various individual traits is comparatively uncertain.

V

THE CONSISTENCY OF TRAIT-RATINGS, IN RELATION TO THE SIGNIFICANCE OF THE TRAIT RATED AND THE CONFIDENCE OF RATING

A. CONSISTENCY IN RELATION TO THE SIGNIFICANCE OF THE TRAIT RATED

The California Behavior Inventory for Nursery School Children consists of descriptions of 231 categories of behavior, termed "traits." It is too much to hope that any one person would, for any considerable number of children, be able to make perfectly valid ratings on each of this large number of traits. We should, however, hope that ratings on traits which are of special significance in a child's personality would be rated with more accuracy than traits of less significance. In the present study, the significance of a particular trait for a particular child is indicated by the number of teachers who starred their first rating of the trait for that child.[18] The problem of the present section is to determine whether or not starred ratings are better judged than unstarred. This will be

[18]The critical reader will question the validity of *this method* of classifying the significance of traits. We do not ourselves suppose that each individual teacher's judgment is free from error in this connection. We do, however, consider that the starred ratings for a particular child apply (on the average) to traits of greater significance to that child than the unstarred. The basis of this belief is *in part merely* observational, and in part statistical. Illustrative of the statistical data in support of this belief are the facts on the consistency of starring, already presented in Chapter III,

considered in terms of the consistency of the average ratings from the three teachers; the question of inter-judge agreement for the starred *vs.* unstarred ratings is reserved for treatment at a later time.

It is necessary, first, to define certain terms. An average rating or re-rating for a child on a given trait will be considered to have a "significance score of 0," if *not any* judge starred her *first rating* of this trait for the particular child in question. (It is immaterial to our classification whether or not any teacher starred her re-rating of the trait.) Similarly, an average rating or re-rating will be considered to have a "significance score of 1," if *at least one* of the judges starred her first rating of this trait for the child; a "significance score of 2," if *at least two* of the judges starred their first rating of this trait for the child; and a "significance score of 3," if *at least three* (in our case, all three) of the judges starred their first rating of this trait for the child.[14]

[14]The classification of average ratings wholly on the basis of the judges' first ratings is justified by the fact that in most preschool situations re-ratings are very seldom made. We have wished our classification, so far as possible, to have reference to practical circumstances and to be repeated readily by other investigators.

It will be noticed that the categories of starred trait-ratings are not mutually exclusive. Thus, if both judge A and judge B starred their rating of trait No. 1 for child No. 10, then the average rating of this trait for child No. 10 would be classified as having a significance score of 1, and also a significance score of 2. The purpose of this system is to obtain at least moderately comparable results from the application of the same definitions in later studies, when the number of judges may be somewhat smaller (or greater) than three. If, for example, only *two* judges were available, and both of these had starred their rating of a trait for a child, this trait is probably more similar to those in our class of ratings starred by *at least two* of three judges than to a class of ratings starred by two and *only two* judges.

We require, now, a statistical measure of relation which will permit ready quantitative comparison of results for the four classes of ratings just defined above. The first measure to suggest itself, of course, is the coefficient of correlation; another, the coefficient of mean square contingency; a third, the mean difference (without regard to sign) between average first ratings and average re-ratings; and fourth, the mean difference without regard to sign, *in relation to the mean difference* (without regard to sign) *which might be expected to occur purely by chance.* All of these statistical measures have been applied to our data, with detailed results as given in Appendix C. The present chapter, however, has been written in terms of the fourth measure. This measure has seemed to us not only the simplest and most direct, but also on theoretical grounds the most justifiable for present comparative purposes (cf. Appendix C).

The use of the mean difference without regard to sign, in relation to the mean difference which might be expected to occur purely by chance, involves the computation of two statistical constants: (*a*) $MD_{aa'}$—the mean difference without regard to sign between the average first rating by three judges and the average re-rating; and (*b*) the chance mean difference without regard to sign, here symbolized by $CMD_{aa'}$. The chance mean difference between average ratings and re-ratings may be computed by substituting for the actual cell-frequencies of the correlation chart, the "independence values" or frequencies to be expected by chance (3, p. 198), and then proceeding exactly the same as in the computation of the $MD_{aa'}$. The fraction $\frac{MD_{aa'}}{CMD_{aa'}}$ represents the proportion which the obtained mean difference is of the chance mean difference. The proportion of agreement *beyond what might be expected by chance* is repre-

TABLE 13

CONSISTENCY OF TOTAL AVERAGE RATINGS

Average First Rating

Average Re-rating	1	1.3	1.7	2	2.3	2.7	3	3.3	3.7	4	4.3	4.7	5	5.3	5.7	6	6.3	6.7	7	Total
7																			1	1
6.7														1		1	1	1	3	7
6.3															1		4	2		7
6												2			4	4	4	1		15
5.7							1				1	1		5	1	5	4	2		20
5.3											1	7	9	5	6	2	3			33
5										1	5	18	16	7	7	5	4			63
4.7								5	4	14	20	26	29	6	4	1				109
4.3							2	7	14	23	34	20	11	5	1					117
4				1		4	6	25	20	44	35	19	5		1					160
3.7					6	7	17	26	38	20	9	3	1	1	1					129
3.3				2	8	6	22	28	24	16	6	2	1							115
3				4	5	12	24	16	7	2	2	1								73
2.7				2	9	10	4	1		1		1								28
2.3		1	2	3	3	3	3					1								16
2			8	5	2	2	3													20
1.7	3	1	1	3	2															10
1.3	2		2	1																5
1	2																			2
Total	7	2	13	21	35	44	82	108	107	121	113	101	72	30	26	18	20	6	4	930

sented by the quantity $1 - \frac{MD_{aa'}}{CMD_{aa'}}$. To obtain the *percentage* of agreement beyond chance, one need merely multiply this quantity by 100; viz., $100 \left(1 - \frac{MD_{aa'}}{CMD_{aa'}}\right)$.

In Table 13, we give a complete representation of the relation between *all* the average first ratings of the three teachers and the average re-ratings. This table may serve more or less as our "standard of comparison." In this table, the percentage of agreement beyond chance is 62.

The percentage of agreement beyond chance for average ratings and re-ratings in the various significance-classes is given in Table 14. In this table a

TABLE 14

PERCENTAGE OF AGREEMENT BEYOND CHANCE FOR AVERAGE RATINGS IN VARIOUS SIGNIFICANCE-CLASSES

Class	n	% agreement beyond chance
All average ratings	910	62
Significance score of 0	587	48
" " " 1	225	61
" " " 2	86	71
" " " 3	32	80

steady rise in agreement is observed as one proceeds from the lower to the higher significance-classes. Tables 15-18 give, in condensed form, the correlation charts from which the data of Table 14 were computed.

TABLE 15

CONSISTENCY OF AVERAGE RATINGS WITH A "SIGNIFICANCE SCORE" OF 0

Average First Rating

Av. Re-rating	1*	2	3	4	5	6	7	Total
7*							1	1
6			1	1	3	3		8
5			3	39	77	7		126
4			55	221	45	1		322
3			70	51	3			124
2			5		1			6
1								. . .
Total	. . .	. . .	134	312	129	11	1	587

TABLE 16

CONSISTENCY OF AVERAGE RATINGS WITH A "SIGNIFICANCE SCORE" OF 1

Average First Rating

Av. Re-rating	1*	2	3	4	5	6	7	Total
7*								. . .
6					3	6	1	10
5			2	5	44	11		62
4		5	37	16	19			77
3		11	46	7	2			66
2		5	5					10
1								. . .
Total	. . .	21	90	28	68	17	1	225

*In these tables the marginal index of "1" stands for 1 and 1 1/3; of "7," for 6 2/3 and 7; of "2," for 1 2/3, 2, and 2 2/3; etc.

TABLE 17

CONSISTENCY OF AVERAGE RATINGS WITH A "SIGNIFICANCE SCORE" OF 2

Average First Rating

Av. Re-rating	1*	2	3	4	5	6	7	Total
7*						2	1	3
6					2	11	2	15
5				1	2	10		13
4		2	2			2		6
3		18	7					25
2	1	19	1					21
1		3						3
Total	1	42	10	1	4	25	3	86

TABLE 18

CONSISTENCY OF AVERAGE RATINGS WITH A "SIGNIFICANCE SCORE" OF 3

Average First Rating

Av. Re-rating	1*	2	3	4	5	6	7	Total
7*					1		3	4
6						7	2	9
5						4		4
4					1			1
3		1						1
2	4	5						9
1	4							4
Total	8	6	. . .	. . .	2	11	5	32

*See footnote to Tables 15 and 16.

B. Consistency in Relation to the Confidence of Rating

Here we consider, on much the same lines as in the previous section, the consistency of average trait-ratings in relation to the confidence of rating. The average ratings and re-ratings are, as before, the average of the three nursery-school teachers' judgments; the average confidence-rating is, however, based on the confidence-ratings of two judges only, viz., judges A and B.[16] The basis for classification, both of average ratings and re-ratings, is the average degree of confidence recorded at the *first* rating, without regard to the degree of confidence expressed at the re-rating. Thus, an average rating or re-rating for a child on a given trait is considered to have been made with "high confidence" when *both* judge A's and judge B's first confidence-ratings were high; "above-average confidence," when one of the judge's confidence-ratings was high and the other's average; etc. (cf. p. 9).

Table 19 below presents the percentage of agreement

TABLE 19

Percentage of Agreement beyond Chance for Average Ratings in Various Confidence-Classes

Class	*n*	Percentage agreement beyond chance
All average ratings	930	62
Low confidence	40	58
Below-average confidence	147	48
Average confidence	362	53
Above-average confidence	242	58
High confidence	139	75

[16]See footnote 9, page 42.

beyond chance for each of the five classes of average ratings which have been distinguished on the basis of A's and B's confidence-ratings. Table 19 shows that except for the small number of ratings in the class of low confidence, a steady (though comparatively small) increase in consistency is noted as one proceeds from the classes of lower to those of higher confidence. The consistency of ratings in the high-confidence class is particularly superior.

The correlation charts from which the data of Table 19 were computed are presented in condensed form in Tables 20-24.

TABLE 20

CONSISTENCY OF AVERAGE RATINGS MADE WITH HIGH CONFIDENCE AT FIRST RATING

Av. Re-rating \ Average First Rating	1*	2	3	4	5	6	7	Total
7*					1	2	4	7
6					3	16	4	23
5			2		9	13		24
4		2	9	6	4	1		22
3		16	11	1				28
2	5	23	1					29
1	4	2						6
Total	9	43	23	7	17	32	8	139

*See footnote to Tables 15 and 16.

TABLE 21

CONSISTENCY OF AVERAGE RATINGS MADE WITH ABOVE-AVERAGE CONFIDENCE AT FIRST RATING

Average First Rating

Av. Re-rating	1*	2	3	4	5	6	7	Total
7*								...
6			1	1	2	4		8
5			1	11	46	9		67
4		2	35	35	15	1		88
3		10	41	14	1			66
2		6	7					13
1								...
Total	...	18	85	61	64	14	...	242

TABLE 22

CONSISTENCY OF AVERAGE RATINGS MADE WITH AVERAGE *CONFIDENCE AT FIRST RATING*

Average First Rating

Av. Re-rating	1*	2	3	4	5	6	7	Total
7*								...
6						6	1	7
5			1	20	42	3		66
4		1	37	131	28			197
3		2	50	34	3			89
2			2					2
1		1						1
Total	...	4	90	185	73	9	1	362

*See footnote to Tables 15 and 16.

TABLE 23

CONSISTENCY OF AVERAGE RATINGS MADE WITH BELOW-AVERAGE CONFIDENCE AT FIRST RATING

Av. Re-rating \ Average First Rating	1*	2	3	4	5	6	7	Total
7*							1	1
6					3	1		4
5				10	15	7		32
4		2	11	54	17	1		85
3		2	14	6	1			23
2			1		1			2
1								. . .
Total	. . .	4	26	70	37	9	1	147

TABLE 24

CONSISTENCY OF AVERAGE RATINGS MADE WITH LOW CONFIDENCE AT FIRST RATING

Av. Re-rating \ Average First Rating	1*	2	3	4	5	6	7	Total
7*								. . .
6								. . .
5			1	4	11			16
4			2	11	1			14
3			7	3				10
2								. . .
1								. . .
Total	. . .	. . .	10	18	12	. . .	. . .	40

*See footnote to Tables 15 and 16.

C. Consistency in Relation to Both the Significance of the Trait Rated and the Confidence of Rating

The 930 average ratings and re-ratings provide sufficient material for a two-way classification—first, on *the basis of confidence of rating*; and, *second, on the* basis of the estimated significance of the trait for the child being rated. Thus, there are 362 average ratings and re-ratings which belong in the average-confidence class; of these, 307 belong in the class having a significance score of 0, 46 in the class with significance score of 1, and 9 in the class with significance score of 2. The consistency for each confidence-significance class in which n is 30 or more is presented in Table 25. This table shows that, for ratings in the

TABLE 25

Consistency of Average Trait-Ratings in Relation to Both the Significance of the Trait Rated and the Confidence of Rating

Class	n	Percentage agreement beyond chance
All average ratings	930	62
Significance score 0:		
Low confidence	36	55
Below-average confidence	122	44
Average confidence	307	46
Above-average confidence	109	50
Significance score 1:		
Average confidence	49	62
Above-average confidence	113	60
High confidence	39	65
Significance score 2:		
High confidence	59	72

same confidence-class, the agreement of judges increases significantly as the estimated significance of the trait for the child being rated increases. For ratings in the same significance-class, however, a similar rise with increase in confidence is not clear. Table 25 encourages the belief that the greater consistency of high-confidence ratings in Table 19 is mediated through the relation between confidence of rating and significance score.

SUMMARY

Average ratings and re-ratings by the three nursery-school teachers were classified into four groups on the basis of the significance of each particular trait in the personality of the particular child being rated. An average rating or re-rating was considered to have a "significance score" of 0, 1, 2, or 3, according as *none*, *at least one*, *at least two*, or *all three* of the nursery-school teachers starred their first ratings of the particular trait for the particular child. A steady improvement is observed in the consistency between average ratings and re-ratings as one proceeds from the unstarred to the multiply starred classes. The percentage of agreement beyond chance for the average ratings in class 0 is 48; in class 1, 61; in class 2, 71; and in class 3, 80.

A similar method of investigation was pursued to study the relation between consistency of average ratings and the confidence of rating. Each average rating and re-rating for a child on a given trait was classified into one of five groups, on the basis of the degree

of confidence recorded by the teachers at their first rating of the trait for the child. It was found that ratings in the very highest confidence category were made with the most consistency (percentage of agreement beyond chance, 75); ratings in the other confidence categories, however, showed only slight differentiation with respect to consistency (average percentage of agreement beyond chance, about 55). Further analysis indicates that the greater consistency of ratings made with higher confidence is probably mediated through the relation between confidence of rating and significance score.

Investigators of the reliability and validity of personality ratings should recognize that in general, when teachers are asked to rate individual children on a long list of traits or attributes, the traits which are of little apparent significance in a child's personality will, at least for that child, have been but casually noticed; such traits of presumably lesser significance will probably be only poorly judged. Traits considered by a judge to be of greater importance are, in general, rated with definitely superior consistency. The analysis in the present chapter seems to have justified the requirement that judges indicate (by a star after their ratings) those traits which they consider of special importance in the personality of the child being rated.

VI

SUMMARY AND CONCLUSIONS

The Data. The data of the present study are based primarily on three nursery-school teachers' ratings and re-ratings of 30 children on 31 traits from the *California Behavior Inventory for Nursery School Children*. Both ratings and re-ratings were made on a seven-point scale, in accordance with detailed directions; the re-ratings were made from one to two months after the children had been rated on the total 231 traits of the complete Behavior Inventory. Besides the trait-ratings, each teacher also indicated the degree of confidence with which she made every rating; and in addition indicated (by a star after her rating) whether she considered the particular trait being rated of special significance in the personality of the child being rated. The present study is concerned primarily with the consistency of ratings, or self-agreement of a judge from one occasion to the next; this is studied in relation to the trait being rated, the child being rated, the confidence of rating, and the estimated importance of the trait for the child being rated. The consistency of the confidence-ratings is also considered, as well as the consistency of "starring." Inter-judge agreement in trait-ratings is also briefly treated.

Consistency of Ratings of Traits. As measured by the coefficient of correlation, the consistency of a single judge's ratings of a trait is, on the average, about .70; if the ratings of three teachers are pooled, the consis-

tency correlation averages about .85. Wide differences are found in results for the various traits.

Consistency of Ratings of Children. Consistency correlations, with the *individual child* as the statistical unit, may be obtained by correlating the ratings and re-ratings on the various traits for the particular child in question. The consistency of a single judge's ratings of a child averages about .65; this rises to .80 if ratings by three teachers are pooled. As in the case of the individual traits, the consistency correlations for the individual children vary widely.

Inter-Judge Agreement in Ratings. The agreement, in terms of correlation, between one judge's ratings of a trait and another's averages somewhat over .55; the agreement between the pooled ratings of three teachers and an infinite number of other teachers is estimated, on the average, as over .80. The r's average slightly lower than this when the individual child, instead of the individual trait, is the statistical unit.

Consistency of "Starring." With respect to the starring of trait-ratings, the consistency of a single teacher's starring of individual trait-ratings is too low to encourage confidence. The pooled judgment of the three teachers is, however, more satisfactory, being equal to a coefficient of contingency (from a 4 x 4 table) of .55. (The highest possible C from a 4 x 4 table is .866.)

Each teacher's evaluation of the general (or group) importance of each trait was obtained simply by counting the total number of times she starred her rating of the trait. This number was correlated with the

number of times she starred her re-ratings of the trait. The correlation, for the individual teacher, averages no higher than around .50; for the pooled judgment of three teachers, the consistency correlation is slightly over .70. A companion problem is the consistency of the total number of trait-ratings starred for each individual *child*. The consistency correlations for this are much higher than for the individual traits—being around .79 for an individual teacher, and .89 when the judgments of three teachers are pooled. Evidently the number of starred ratings per child remains more or less constant from rating to re-rating; less certainty exists, however, as to the particular traits whose ratings are to be starred.

Consistency of Confidence-Ratings. The judges' confidence-ratings were studied in much the same way as the starring of trait-ratings. For the consistency of the average confidence with which two judges rate and re-rate all the children on all the traits ($n=930$), the coefficient of contingency is about .50. The consistency of the average confidence with which two teachers rate a single trait for all children is expressed by a coefficient of correlation of about .45; this is surprisingly low, in view of the contingency coefficient just cited above. If, however, the confidence-ratings for a given child on all the traits are averaged, the consistency found is moderately high (r=about .80). It appears, then, that in general the judges are moderately consistent in the average confidence of their ratings for children; but the distribution of their confidence among the various individual traits is comparatively uncertain.

Consistency of Ratings in Relation to the Confidence of Rating and the Estimated Importance of the Trait for the Child Being Rated. The final, and statistically most difficult problem, was to study the consistency of the average ratings of the three judges, in relation to the average confidence of rating, and the starring of trait-ratings. With respect to the latter, each average rating or re-rating was assigned a "significance score" of 0, 1, 2, or 3, according as *none, at least one, at least two,* or *all three* of the nursery-school teachers starred their *first* ratings of the particular trait for the particular child. The consistency of the ratings in each of these classes was then studied separately. Between average ratings and re-ratings in the significance-class of 0 the percentage of agreement beyond chance is 48; in class 1, the percentage of agreement beyond chance is 61; in class 2, 71; and in class 3, 80.

A similar method of investigation was used to study the relation between consistency of average ratings and the confidence of rating. It was found that ratings in the very highest confidence category are made with the most consistency (percentage of agreement beyond chance, 75); ratings in the other confidence categories, however, show only slight differentiation with respect to consistency (average percentage of agreement beyond chance, about 55). Further analysis indicates that the greater consistency of ratings made with higher confidence is probably mediated through the relation between confidence of rating and significance score.

Results from Other Measures of Relation. Various

other measures of relation, besides the "percentage of agreement beyond chance," were employed in studying the relation between consistency of rating, confidence, and starring. The coefficient of correlation and the coefficient of contingency yield results qualitatively similar to the percentage of agreement beyond chance; the mean difference without regard to sign, however, gives results which in general run counter to those from r, C, and the percentage agreement beyond chance. This is undoubtedly due to the fact that the mean difference without regard to sign fails to take account of such agreement as might be due to chance.

Spurious Factors. Statistical examination of the possible influence of various spurious factors indicates that such influence is either slight or absent. The results obtained in the present study for the 31 traits from the California Behavior Inventory are shown to be representative of what may be expected, on the average, from the total 231 traits of the complete Inventory.

Interpretation. The facts given above are susceptible to varying interpretations; we shall attempt to be very brief. The individual differences observed among the various traits and children seem fundamental; the causes of these individual differences, unfortunately, seem to be largely unknown. In spite of the fact that ratings (in one form or another) are used regularly in connection with matters of the highest personal, professional, and political significance, psychology seems to have regarded the problems of rating as of merely incidental importance. This atti-

tude is the less understandable in view of such results as have been outlined above. Certainly the pooled ratings of competent judges possess sufficient consistency to permit active research in this field.

The objection may be raised that the results of the present study were obtained in a sample of nursery-school children; such a sample perhaps offers certain advantages to the judge, from the point of view of ease of observation and simplicity of behavior (*2a*). Even if this were the case, however, it could hardly justify the comparative neglect of problems of rating at the preschool level.

The first requirement for more rapid progress in the field of ratings is a less prejudiced attitude among psychologists toward the topic of rating. The second requirement is the training of good judges. The third requirement is the pursuit of detailed investigations whereby the fund of reliable information concerning ratings will be improved. Such studies should include not only ratings of average behavior, but also ratings of the variability of behavior; they should include not merely three judges observing the subjects under the same environmental conditions, but many more judges observing under varying conditions; they should include not merely studies of children as of a certain particular date, but also follow-up studies which extend throughout a considerable portion of each subject's total life-span. These are, of course, large tasks; but if, as now seems apparent, ratings must remain a fundamental source of information concerning personality, such undertakings cannot be avoided without risking a serious deficiency in our supply of research techniques.

APPENDIX A

DISTRIBUTIONS OF RATINGS

Within the scope of this brief report, it is not possible to present for each trait or child separately the frequency distribution of each judge's ratings. Tables 1-4 of Chapter II do, however, present information concerning the standard deviations of the trait-ratings for each individual child and trait.[10] In Table 26 we present the frequency distributions of each judge's ratings and re-ratings for all children and traits together. The other tables, given in the text, will supply the critical reader with much other basic material which it is not practical to repeat in this place.

TABLE 26

DISTRIBUTIONS OF RATINGS OF THIRTY SUBJECTS ON THIRTY-ONE TRAITS

	Judge A		Judge B		Judge C	
Rating	First rating n	Re-rating n	First rating n	Re-rating n	First rating n	Re-rating n
7	43	5	12	10	36	16
6	95	46	64	52	79	70
5	186	199	175	192	192	242
4	266	437	371	395	271	277
3	214	190	223	209	217	217
2	94	47	71	56	100	91
1	32	6	14	16	31	17
Total	930	930	930	930	926	930
Mean	4.01	4.00	3.93	3.95	3.94	3.98
Median	3.97	4.01	3.92	3.97	3.92	4.01
S.D.	1.40	.962	1.12	1.07	1.37	1.23

APPENDIX B

THE REPRESENTATIVENESS OF THE THIRTY-ONE SELECTED TRAITS

The California Behavior Inventory for Nursery-School Children

[10] The means are not included in Tables 1-4, principally because these in general varied but little from the midpoint of the seven-point scale (namely, 4).

lists descriptions of 231 traits; the present study has been restricted to 31 of these. Are the results for these 31 traits similar, in general, to the results which may be expected from the other traits of the Inventory?

A direct answer to this question would require that the nursery-school teachers not only rate each of our 30 children on all 231 traits of the Inventory (which they have done), but also re-rate them. Such an extensive undertaking was not feasible. We have, however, compiled considerable evidence which shows that the 31 traits do yield results which are, in general, reasonably representative. This evidence consists of comparisons of results obtained from the first rating of the 31 traits with results obtained from the first rating of all the traits of the Inventory. Tables 27-29 present the facts. In Tables 27-28 it can be seen that the 31 traits are similar

TABLE 27

DISTRIBUTIONS OF RATINGS OF (*a*) 31 SELECTED TRAITS AND (*b*) TOTAL 231 TRAITS OF THE INVENTORY

Judge	Sample	Percentage frequency of the following ratings: 1	2	3	4	5	6	7
A	31 traits	3.4	10.1	23.0	28.6	20.0	10.2	4.6
A	Total Inventory	3.3	9.8	20.7	31.7	20.7	11.1	2.7
B	31 traits	1.5	7.6	24.0	39.9	18.8	6.9	1.3
B	Total Inventory	1.0	6.1	22.9	40.8	20.9	6.7	1.7
C	31 traits	3.3	10.8	23.4	29.3	20.7	8.5	3.9
C	Total Inventory	2.1	8.5	23.8	35.4	21.1	7.1	2.1

TABLE 28

DISTRIBUTION OF CONFIDENCE-RATINGS ACCOMPANYING THE TRAIT-RATINGS OF (*a*) 31 SELECTED TRAITS AND (*b*) TOTAL 231 TRAITS OF THE INVENTORY

Judge	Sample	Percentage frequency of the following ratings: Low	Average	High
A	31 traits	18.2	49.9	31.9
A	Total Inventory	18.3	46.9	34.8
B	31 traits	13.4	67.3	19.3
B	Total Inventory	16.2	65.3	18.5
C	31 traits	41.7	46.5	11.8
C	Total Inventory	51.0	45.9	3.1

TABLE 29

CORRELATION BETWEEN STATISTICAL CONSTANTS FOR EACH CHILD BASED ON RATINGS OF (*a*) 31 SELECTED TRAITS AND (*b*) TOTAL 231 TRAITS OF THE INVENTORY

Variables correlated	r	Mean (31)	Mean (231)
r_{AB} (31) *and* r_{AB} (231)	.83*	.45†	.47‡
r_{AC} (31) *and* r_{AC} (231)	.82	.44	.43
r_{BC} (31) *and* r_{BC} (231)	.80	.48	.44
MD_{AB} (31) *and* MD_{AB} (231)	.80	.90	.86
MD_{AC} (31) *and* MD_{AC} (231)	.86	1.03	.95
MD_{BC} (31) *and* MD_{BC} (231)	.75	.91	.84
σ_A (31) *and* σ_A (231)	.986	1.32	1.26
σ_B (31) *and* σ_B (231)	.95	1.01	1.02
σ_C (31) *and* σ_C (231)	.95	1.30	1.15
Starred ratings, A (31) *and* starred ratings, A (231)	.96§	24.9%‖	21.0%‖
Starred ratings, B (31) *and* starred ratings, B (231)	.95	13.7%	9.0%
Starred ratings, C (31) *and* starred ratings, C (231)	.88	14.5%	11.2%
Av. confidence, A (31) *and* av. confidence, A (31)	.95	2.20	2.17
Av. confidence, B (31) *and* av. confidence, B (31)**	.95	2.12	2.03

*To be read: The r between ratings by judge A and judge B for each child on 31 traits is correlated with the r between ratings by judge A and judge B for each child on 231 traits to the extent of .83.

†To be read: The mean correlation between ratings by A and B on each child for 31 traits. Cf. Table 4.

‡To be read: The mean correlation between ratings by A and B on each child for 231 traits.

§To be read: The no. of starred ratings by judge A for each child on 31 traits is correlated with the no. of starred ratings by judge A for each child on 231 traits to the extent of .96.

‖To be read: The percentage of judge A's ratings of 30 children on 31 traits which was starred is 24.9; the corresponding percentage for A's ratings of 30 children on 231 traits is 21.0.

**No correlation is available between judge C's confidence-ratings for each child, since C did not discriminate between children when making her confidence-ratings (cf. p. 42).

to the total 231, with respect to the distribution of trait-ratings, and to the confidence with which trait-ratings were made. Statistical constants for each individual child, based on ratings of the 31 traits, show a fairly high correlation with constants based on ratings of the total 231; moreover, the means of these constants, for the 31 and the total 231 traits, are rather closely similar (Table 29). Tables 27-29 indicate rather clearly that the results obtained for the *31* traits may, *with only slight error, be accepted as reasonably* representative of what should be expected, on the average, from the full 231 traits of the Behavior Inventory.

APPENDIX C

THE CONSISTENCY OF TRAIT-RATINGS IN RELATION TO THE SIGNIFICANCE OF THE TRAIT RATED AND THE CONFIDENCE OF RATING

A. Measures of Relation

In Chapter V extensive use was made of a measure of relation termed the "percentage of agreement beyond chance." It remains to consider the relative suitability of other statistical measures and the results which are obtained from their use.

The *coefficient of correlation* ($r_{BB'}$) gives results which are, in general, qualitatively similar to those obtained from the use of the percentage agreement beyond chance. The rise in correlation from one class of ratings to another is, however, comparatively exaggerated, due to the sensitiveness of r to extreme cases (cf. Tables 15-18, 20-24, and 30-32).

The *coefficient of contingency* ($C_{BB'}$) also tends in general to give results qualitatively similar to the percentage agreement beyond chance. The contingency coefficient, however, shows a comparatively dampened rise from one class of ratings to another (cf. Tables 30-32). Interpretation of the results for C should bear in mind that C is best adapted for the study of relations between qualitative categories, rather than quantitative variables. As a consequence, C makes no direct distinction between a quantitatively small disagreement and a quantitatively large one; also, C makes no direct distinction between agreement at the extremes of the distributions and agreement in the middle. The doubtful sensitivity of C to

these quantitative aspects tends to diminish its usefulness in the comparative study of the relations in Tables 15-18 and 20-24.

The coefficient of contingency has frequently in the past been carelessly or uncritically used. We have never, ourselves, computed C for any contingency table unless n were greater than 100. All the coefficients of contingency reported in the present Appendix were computed from 7 x 7 tables (except the C of .64 in Table 32, computed from a 6 x 6 table). The particular grouping used in each contingency table was determined mainly by the requirement that the chance frequency in any cell should never be less than 1 (5, p. 265). Re-ratings were always grouped the same as the corresponding first ratings.[17] In comparing results by C with those from other measures, it should be remembered that the highest possible C in a symmetrical 7 x 7 table is .926 (7, p. 66).

The *mean difference without regard to sign* ($MD_{\text{ab}'}$) is the simplest measure which we have used. It gives results which, in general, run counter to those given by r, C, and the percentage agreement beyond chance. This is undoubtedly due to the fact that the MD does not take account of such agreement as might be due to chance. The MD tends to be small when the standard deviations of the ratings are small, and the agreement which might be due to chance is high (cf. Tables 30-32).

The *percentage agreement beyond chance* (defined on p. 52) is the measure of relation which has seemed to us, on the whole, the most acceptable. To be sure, this measure, like all others, is open to objection. The chief criticism probably is the fact that, given a correlation chart with homoscedastic arrays (and hence equal "standard errors of measurement" throughout the chart), the percentage of agreement beyond chance will be larger for arrays at the extremes of the correlation chart than for any group of arrays at the center. (This is due to the fact that the chance mean difference is greater for the arrays at the extremes.) But this criticism, in reality, is nothing but an argument in favor of the mean difference without regard to sign (MD)—since no conventional measure

[17]The application of these rules usually left little choice concerning the groups to be used in a contingency table, unless n were quite large. Table 13 (in which n=930) was reduced to a 7 x 7 basis by the use of the following groups: 1-2.7, 3-3.3, 3.7, 4, 4.3, 4.7-5, and 5.3-7.

of relation except the *MD* is altogether exempt from this criticism. The *MD* is hardly acceptable, however, unless one is willing to assume that all agreement is *bona fide* and not due in any part to chance. The more conventional and conservative statistical assumption is exactly the opposite, viz., to accept no agreement as *bona fide* unless surely *not* due to chance. Possibly neither of these extreme assumptions is the best that could be made; of the two, however, the conventional assumption appears to us definitely the safer.

TABLE 30

CONSISTENCY OF AVERAGE RATINGS IN VARIOUS SIGNIFICANCE-CLASSES*

Class	n	% agreement beyond chance	$r_{33'}$	$C_{33'}$	$MD_{33'}$	σ_3	$\sigma_{3'}$
All average ratings	930	62	.81	.75	.42	1.12	.94
Significance score 0	587	48	.68	.65	.37	.65	.65
Significance score 1	225	61	.85	.72	.48	1.14	.96
Significance score 2	86	71	.92	..†	.59	1.89	1.58
Significance score 3	32	80	.97	..†	.51	2.50	2.13

*For definition of the various classes of ratings and of the symbols of the table, cf. Chapters II and V. The symbols σ_3 and $\sigma_{3'}$ refer to the standard deviation of average first ratings and average re-ratings, respectively; these have been included in the table to facilitate interpretation of the measures of relation.

†Not computed because of the small value of n.

TABLE 31

CONSISTENCY OF AVERAGE RATINGS IN VARIOUS CONFIDENCE-CLASSES*

Class	n	% agreement beyond chance	$r_{33'}$	$C_{33'}$	$MD_{33'}$	σ_3	$\sigma_{3'}$
All average ratings	930	62	.81	.75	.42	1.12	.94
Low confidence	40	58	.80	..	.34	.74	.66
Below-av. confidence	147	48	.71	.70	.44	.83	.70
Average confidence	362	51	.75	.69	.36	.71	.66
Above-av. confidence	242	58	.80	.70	.47	1.02	.91
High confidence	139	75	.94	.80	.51	1.88	1.61

*The footnotes to Table 30 apply also to the present table.

TABLE 32

CONSISTENCY OF AVERAGE TRAIT-RATINGS IN RELATION TO BOTH THE SIGNIFICANCE OF THE TRAIT RATED AND THE CONFIDENCE OF RATING*

Class	n	% agreement beyond chance	$r_{\mathrm{I\,II}}$	$C_{\mathrm{I\,II}}$	$MD_{\mathrm{I\,II}}$	σ_{I}	σ_{II}
All average ratings	930	62	.81	.75	.42	1.12	.94
Low confidence:							
Significance score 0	36	55	.77	..	.34	.71	.61
Below-average confidence:							
Significance score 0	122	44	.65	.64	.43	.74	.67
Average confidence:							
Significance score 0	307	46	.67	.66	.35	.57	.56
Significance score 1	46	62	.83	..	.44	1.04	.87
Above-average confidence:							
Significance score 0	109	50	.69	.70	.41	.66	.77
Significance score 1	113	60	.84	.71	.48	1.11	.95
High confidence:							
Significance score 1	39	65	.88	..	.50	1.29	1.15
Significanc score 2	59	72	.94	..	.58	1.99	1.66

*The footnotes to Table 30 apply also to the present table.

B. SPURIOUS FACTORS

We have not yet considered the possibility that spurious factors may be responsible for the apparent increase in the consistency of rating with increase in estimated importance of the trait.

1. Does a differential memory factor operate to produce the differences observed between the classes of ratings in Table 14? This seems most unlikely when one recalls that the *nursery-school* teachers did not begin to re-rate until one to two months after completing the first ratings. Inasmuch as the *first ratings* numbered around 7000 per teacher (30 children rated on 231 traits), the memory factor in re-rating would, at least on *a priori* grounds, seem to be rather uniformly negligible.

2. Does a differential age factor operate to produce the differences observed between the classes of ratings in Table 14? Suppose, for example, that the trait-ratings of older children were starred much more often than those of younger. Starred trait-ratings tend

to be rather extreme—generally 2 or 6, practically never 4. Because of this association between starring and trait-rating, it is possible that the higher consistency of the starred trait-ratings reflects only a consistency in starring; and the consistency of starring itself may be spurious, if strongly related to the age of the children being rated. Actually, however, the basic assumption of this whole argument falls down, since the correlation between chronological age and the number of trait-ratings starred for each child is roughly zero (or slightly negative).

3. Finally, one may be tempted to inquire whether the measures of consistency of rating increase as they do with increase of estimated trait-importance, because of statistical manipulation. We believe the answer to this is negative. It should be stressed, in this connection, that all the average ratings and re-ratings were classified *wholly* on the basis of starring at the *first* rating. Had the assignment of a significance score of 3, for example, required that all three teachers have starred both their first ratings *and* their re-ratings of a trait, then it might be argued that the agreement between average trait-ratings with a significance score of 3 is only a function of the enforced agreement in starring at the first and second rating. But our classification imposed *no restriction at all* on the starring of the *re*-ratings.

A similar investigation of possible spurious influences on the results in Table 19 (relating to confidence of judgment) fails to discover any significant factor. The conclusion appears to stand, then, that the consistency of average trait-ratings increases with increase in the confidence of judgment, and especially with increase in the estimated importance of the trait for the child being rated.

References

1. CHESIRE, L., SAFFIR, M., & THURSTONE, L. L. Computing diagrams for the tetrachoric correlation coefficient. Chicago: Univ. Chicago Press, 1933.

2. CONRAD, H. S. The California Behavior Inventory for Nursery School Children. Berkeley, Calif.: Univ. Calif. Press, 1933. Pp. 74.

2a. ———. The validity of personality ratings of preschool children. *J. Educ. Psychol.*, 1932, 23, 671-680.

3. GARRETT, H. E. Statistics in psychology and education. New York: Longmans, Green, 1926. Pp. xiii+317.

4. HOLZINGER, K. J. Statistical tables for students in education and psychology. Chicago: Univ. Chicago Press, 1925. Pp. v+74.

5. KELLEY, T. L. Statistical method. New York: Macmillan, 1924. Pp. xi+390.

6. MINER, J. R. Tables of $\sqrt{1-r^2}$ and $1-r^2$ for use in partial correlation and in trigonometry. Baltimore: Johns Hopkins Press, 1922. Pp. 49.

7. YULE, G. U. An introduction to the theory of statistics. (7th ed.) London: Griffin, 1924. Pp. xv+424.

UNE ÉTUDE STATISTIQUE DES ÉVALUATIONS SUR L'ÉCHELLE DE COMPORTEMENT DE CALIFORNIE POUR LES ÉLÈVES DES ÉCOLES MATERNELLES

(Résumé)

Les données de cette étude sont basées principalement sur les évaluations et les nouvelles évaluations de 30 enfants à l'égard de 31 traits sur l'Echelle de Comportement de Californie pour les Eleves des Ecoles Maternelles par trois maîtresses des écoles maternelles. On a fait et les évaluations et les nouvelles évaluations sur une échelle de sept points, selon des directions détaillées; on a fait les nouvelles évaluations après un intervalle d'un à trois mois. Outre les évaluations des traits, chaque maîtresse a indiqué le degré de confiance qu'elle a senti en faisant chaque évaluation; et aussi (au moyen d'un astérisque après son évaluation) si elle considérait que le trait spécifique en train d'être évalué avait une signifiance spéciale pour la personnalité de l'enfant en question. Dans cette étude il s'agit principalement de la constance des évaluations, ou de l'accord d'un juge avec lui-même de l'une occasion à l'autre; on étudie cela sous le rapport du trait en train d'être évalué, de l'enfant évalué, de la confiance de l'évaluation, et de l'importance estimée du trait pour l'enfant évalué. La constance des évaluations de confiance est aussi considérée, ainsi que la constance des astérisques. On discute brièvement l'accord des juges l'un avec l'autre dans l'évaluation des traits.

Comme mesurée par le coefficient de corrélation, la constance des évaluations moyennes pour un trait par trois maîtresses est d'environ 0,85. La constance des évaluations moyennes pour un enfant pour les 31 traits est d'environ 0,80. Il existe de grandes différences de constance et pour les divers traits et pour les divers enfants.

L'accord, en termes de la corrélation, entre les évaluations d'un juge et celles d'un autre, est en moyenne un peu plus de 0,55. On obtient une corrélation un peu moins élevée quand l'enfant individuel, au lieu du trait individuel, est l'unité statistique.

Comme mesurée par le pourcentage de l'accord non expliqué par le hasard, la constance des évaluations à l'astérisque est définiment plus élevée que celle des évaluations sans astérisque. De même manière, les évaluations faites avec une grande confiance sont aussi plus constantes.

On a étudié les facteurs faux mais ils ont exercé une petite influence ou bien ont été absents.

On appuie sur le fait que les évaluations des juges compétents prises ensemble posèdent assez de constance pour permettre une recherche active. A cause de l'usage étendu des évaluations personnelles dans la vie quotidienne, une étude méthodologique plus étendue semble clairement désirable.

CONRAD

EINE STATISTICHE UNTERSUCHUNG AN TESTBEFUNDEN, AM DEM CALIFORNISCHEN EIGENSCHAFTSTESTINVENTORIUM FÜR VOR-SCHULPFLICHTIGE KINDER [CALIFORNIA BEHAVIOR INVENTORY FOR NURSERY SCHOOL CHILDREN] ERMITTELT

(Referat)

Die Befunde der gegenwärtigen Untersuchung gründen sich hauptsächlich auf Rankierungen und Nach-rankierungen [ratings and re-ratings] welche von drei Lehrerinnen aus Kleinkinderschulen [nursery schools] an 30 Kinder in Bezug auf 31 in dem Californischen Eigenschaftsinventorium für vor-schulpflichtige Kinder eingetragenen Karaktereigenschaften ermittelt worden. Sowohl die Rankierungen wie auch die Nach-rankierungen wurden, ausführlichen Anweisungen befolgend, auf Basis einer sieben Abstufungen enthaltenden Massleiter [seven-point scale] berechnet. Die Nachrankierungen wurden ein bis zwei Monaten nach den Rankierungen gemacht. Die Lehrerinnen notierten nicht nur die Rankierungen in Bezug auf die Karaktertereigenschaften [trait-ratings] sondern auch den Grad des Selbstvertrauens mit der jede Rankierung von ihnen gemacht wurde. Ebenfalls gab die Lehrerin an, ob ihrer Meinung nach die augenblicklich in Betracht kommende Eigenschaft in der Persönlichkeit des in Frage kommenden Kindes eine besonders wichtige Rolle spiele. Wenn ja, so wurde der Rankierung ein Stern hinzugefügt. Die gegenwärtige Untersuchung bezieht sich besonders auf die Beständigkeit [consistency] der Rankierungen—d.h., auf die Übereinstimmung der Richterin mit sich selber, bei zwei gleichen Gelegenheiten. Diese Übereinstimmung wurde untersucht in Bezug auf die zu rankierende Karaktereigenschaft, das zu rankierende Kind, das Selbstvertrauen der Lehrerin bei der Rankierung, und die Schätzung der Wichtigkeit der betreffenden Eigenschaft bei dem betreffenden Kind. Es werden u.a. auch die Beständigkeit der Rankierungen in Bezug auf Selbstvertrauen [consistency of the confidence ratings] und die Beständigkeit der "Besternung" ["starring"] berücksichtigt. Der Grad der Übereinstimmung der verschiedenen Richter in Bezug auf die Rankierungen der Karaktereigenschaften wird kurz besprochen.

Nach dem Korrelationskoeffizient gemessen, beträgt die Beständigkeit der durchschnittlichen Rankierungen einer Eigenschaft durch die drei Lehrerinnen ungefähr .85. Die Beständigkeit der durchschnittlichen Rankierungen an einem Kind für alle 31 Eigenschaften beträgt ungefähr .80. Es zeigen sich weite Unterschiede in der Beständigkeit, sowohl zwischen einzelnen Eigenschaften wie auch zwischen einzelnen Kindern.

Die Übereinstimmung, als Korrelationszahl ausgedrückt, zwischen den Rankierungen zweier Richter beträgt etwas mehr als .55. Unter statistischer Berücksichtigung des einzelnen Kindes statt der einzelnen Eigenschaft erhält man eine etwas niedrigere Korrelation.

Es wurde die Beständigkeit der besternten Rankierungen [starred ratings] an dem Prozentsatz gemessen, mit dem der eigentlich erzielte Grad der Übereinstimmung den durch die Einwirkung der Zufalls zu erwartenden überragte. So gemessen erwies sich die Beständigkeit bei besternten Rankierungen bestimmt höher, als bei nicht-besternten. Ähnlicherweise haben auch die mit starkem Selbstvertrauen gemachten Rankierungen eine höhere Beständigkeit als die entgegengesetzten.

Es wurde nach eventuellen spuriösen Einwirkungen geforscht. Diese erwiesen sich aber entweder als abwesend oder als von geringer Bedeutung.

Es wird betont, dass die vereinigten Rankierungen tauglicher Richter genügend beständig sind, um lebhafte Untersuchung zu ermöglichen. Infolge der weit verbreiteten Verwendung der Rankierung der Eigenschaften [personality ratings] im alltäglichen Leben erscheint die weitere methodologische Erforschung der Rankierungen bestimmt wüschenswert.

CONRAD

$7.00 per volume
Single numbers $2.00

MONTHLY
Two volumes per year

August, 1934
Volume XVI, No. 2

GENETIC PSYCHOLOGY MONOGRAPHS

Child Behavior, Animal Behavior, and Comparative Psychology

EDITED BY
CARL MURCHISON

Worcester, Massachusetts

Entered as second-class matter December 1, 1925, at the post-office at Worcester, Mass., under Act of March 3, 1879

$7.00 per volume
Single numbers $2.00
This issue, $3.00

MONTHLY
Two volumes per year

August, 1934
Volume XVI, No. 2

GENETIC PSYCHOLOGY MONOGRAPHS

Child Behavior, Animal Behavior, and Comparative Psychology

AN EYE-MOVEMENT STUDY OF OBJECTIVE EXAMINATION QUESTIONS*

From the Psychological Laboratory of the University of Minnesota

By

ARDEN FRANDSEN

*Accepted for publication by John E. Anderson of the Editorial Board and received in the Editorial Office, July 18, 1933.

Worcester, Massachusetts

Entered as second-class matter December 1, 1925, at the post-office at Worcester, Mass., under Act of March 3, 1879.

NOWLEDGMENTS

The writer is greatly indebted to Dr. Miles A. Tinker, his graduate adviser, for suggesting the problem and for his encouraging and helpful direction and criticism throughout the course of this investigation. He is grateful also to Dr. M. J. Van Wagenen for advice with a statistical problem, and to Dr. W. S. Miller for a suggestion in planning the work and for his kindness in permitting the use of his unpublished Analogies Test. The privilege of discussing certain parts of the problem with Professor D. G. Paterson, Dr. E. G. Williamson, and Dr. S. R. Hathaway was very helpful. Mr. George Haslerud gave able assistance in revising the manuscript.

ARDEN FRANDSEN

UNIVERSITY OF UTAH
SALT LAKE CITY, UTAH

CONTENTS

I

INTRODUCTION

Photographic studies of eye-movements have demonstrated that the progressive stages in learning to read, from the initial attempts of the beginner to the skilful performance of the mature reader, are accompanied by reductions in the number and duration of fixations, by fewer regressive movements, and by an increasingly smooth and regular sequence of saccadic shifts along the lines of print (3). The sequence of eye-movements is further characterized, for the typical type of reading of the mature reader, by one long fixation near the beginning of the line, followed by two or three short ones, and a longer one towards the end of the line (7). This process, however, like most functions which psychologists have studied, is not lacking in variability. The pattern of eye-movements is conditioned by several factors in addition to level of maturity (37, p. 49-96). Eye-movement patterns in reading the vernacular differ considerably according to the purpose for which the material is read. Variations in the sequence, number, and duration of pauses have been found to be dependent on whether the material is read (*a*) for oral reproduction (26, 14, 2); (*b*) to typewrite (6); (*c*) to proof read (12, 7, 37); (*d*) to study (14, 18); and (*e*) to learn to spell (1). Similar variations accompany changes in content. The eye-movements have been noted to differ in response to changes in (*a*) difficulty (12, 18); (*b*) subject-matter (18,

37); (*c*) arrangement of words (19, 37); (*d*) familiarity with the language (19, 5, 18, 20, 17); and (*e*) type of reading symbols including numerals (4, 7, 19, 21, 29, 33), formulae (31, 21), and music (19). These investigations have revealed such extreme variability that it seems doubtful whether the concept of "regular or normal procedure" has any utility except as a point of reference for descriptive purposes; certainly the plural is more appropriate. This means, of course, *that a complete account of the nature of reading*, even for practical purposes, involves a description of reading of the various types of material in all the situations where such reading might be required.

II

PROCEDURE

In this study the aim is twofold: (1) to describe in terms of photographic records of eye-movements how objective examination questions, differing in form and difficulty, are read as they are answered; (2) to make a quantitative analysis of the different types from the point of view of economy of time and possible specific validity. Six types of questions, representing those in most general use (24, p. 189), were chosen for study: (1) wrong-word or cross-out, (2) analogy, (3) multiple-choice, (4) true-false, (5) one-word-completion, and (6) disarranged sentences. The reading of a paragraph of scientific prose was registered in the same manner to serve as a standard of comparison.

Records of the eye-movements of the subjects (66 college sophomores) as they read and answered samples of each of these types were obtained by use of a modification of the Dodge apparatus for photographing eye-movements. No attempt to describe the apparatus is made here since this has already been done very adequately by Dr. Tinker (35, and in greater detail in 31).

From the photographic records thus obtained it is possible to derive six measures of the eye-movements involved in reading: (1) length and duration of eye-movements; (2) number of pauses per line; (3) number of regressions or backward movements per line; (4) duration of pauses; (5) perception time per line; and (6) relative position and sequence of the series of

fixations. Perception time per line is the sum of the pause durations, since, as Dodge has shown (9, 10), there is no clear vision during inter-fixation movements, but to get total reading-time accurately the durations of the eye-movements should be added. In this study, however, since extent and duration of eye-movements constitute only a small fraction of total reading-time, about 5 per cent (32), and because they are very difficult to record, requiring timing to at least 1/100 of a second (34), they have been omitted.

Variables to Control

In order to isolate the differences resulting from only the reading and answering of different types of questions, it was necessary to control the factors considered below:

1. *Content of Material.* The ideal condition from one point of view would have been to use identical content in all the types of questions, but this was precluded by the necessity for satisfying two other conditions: (*a*) If the same content had been used in each set of the different types with a single group of subjects, the effect of practice on succeeding sets would have constituted an uncontrolled variable. (*b*) To have used identical questions for the different types, but with different subjects for each type, would have required a much larger number of subjects and more film than the funds available would have permitted; and even then the possibility of selective factors operating to complicate the results would have existed. The

most adequate compromise appeared to be the use of questions from similar content, but different enough to obviate any possible practice effect on successive items. Consequently the corresponding questions in each type were chosen or composed from somewhat similar general information content. These included items, sometimes modified, from Army Alpha, Otis Self-Administering Tests of Mental Ability, Miller's Analogies Test (not published), and items formulated by the writer.

2. *Relative Difficulty.* A second requirement was the control of possible variation in difficulty of items of the various types. As a means of securing stable measures of relative difficulty, from the point of view of correctness in answering them, 21 questions of each type (105 in all)[1] were given as a "general information test" to 320 sophomores in psychology at the University of Minnesota. The percentage missing each item was recorded. On the basis of this index of relative difficulty, five questions of each type were selected so that corresponding items in each set would be matched approximately and the questions in each set would increase progressively in difficulty. Table 1 shows the degrees of difficulty, stated in terms of percentage missing each item, for the five types of questions.

It is apparent from inspection of this table that a close approximation to equal difficulty for the corresponding items of each type has been achieved, and the averages for each set of five are practically equal.

[1]The disarranged sentence items were not included in this standardization procedure.

(The row of E-averages will be commented upon later.) These values, except for completions, have been corrected for chance because the differential possibilities for chance successes in the various types would have made the raw percentages misleading.

TABLE I

INDEX OF DIFFICULTY: PERCENTAGE (CORRECTED FOR CHANCE) MISSING EACH ITEM

Item	Wrong-word	Analogies	Multiple-choice	True-false	Completion
1	2.9	2.5	2.1	1.9	1.6
2	27.6	25.3	28.4	24.4	26.6
3	37.8	38.8	40.4	38.6	39.1
4	68.7	60.4	53.3	58.2	61.3
5	75.1	82.9	88.4	87.4	79.1
Average	42.4	42.0	42.5	42.1	41.5
E-average	45.5	45.2	40.7	43.3	46.6

The formula used was $\frac{n}{n-1}$, where n equals the number of possible guesses in an item. In true-false questions this is, of course, 2; in the analogy and multiple-choice items it was 4; in the wrong-word questions n may be different for each question, depending upon the number of significant words in the statement; for the completions no correction was applied since there is presumably no limitation to the range of guesses and consequently $\frac{n}{n-1}$ would approach unity.

3. *Printing Arrangement.* A third factor to control, which was easier to deal with but none the less important, was the printing set-up. The items were so

selected or constructed that the variation in length of line would be negligible. The average length was 113 mm. (26.8 picas) with an A.D. for all the items of 3.5 mm. (.84 picas). Ten-point type with 2-point leading was used throughout, and all the material was printed on Suede Finish Book India paper. A paragraph of scientific prose, which was used as a standard of comparison, was also set in exactly the same way, the length of lines being 110 mm. (26 picas). A group of five disarranged sentences in shorter lines, 80 mm. (19 picas), was also included. In each set of five questions the items were arranged in the order of difficulty, from easy to hard, corresponding to the percentages listed in Table 1. They were also printed in the reverse order, hard to easy, with other details remaining exactly the same. Samples of single items from these sets, as well as of the prose, are presented in Figures 1 to 7, pages 110 to 123.

Three numbers in a vertical column were placed at the end of the first line in each set and at the beginning of the last, which the subject was instructed to read, for the purpose of marking points of reference to aid in plotting the records. Similarly, a cross placed to the right of each item, which was fixated after answering the question, served to distinguish the separate items on the film.

4. *Order of Presentation.* In order to equate for any possible advantage to any one type of question from order of presentation, the order was altered systematically so that each set appeared an equal number of times in each position. Also within each set, since

the questions were printed in two forms, one easy to hard (1, 2, 3, 4, 5) and the other hard to easy (5, 4, 3, 2, 1), it was possible, by presenting each form an equal number of times, to control this factor for each question within each set. This, of course, was unnecessary for comparisons according to type but was required only for comparisons according to difficulty.

5. *Selection of Subjects.* As has already been mentioned, the questions were standardized for difficulty on a group of 320 sophomores. To insure the maintenance of the same relative difficulties of the items in the experimental situation it was desirable to have a different group whose performance with the questions would be similar to that of the preliminary standardization group. Therefore, students from similar but not the same classes were chosen for subjects in the experiment proper. Readable records were obtained from 66 sophomores, 33 women and 33 men. Only students who were able to read without spectacles were included. That a very good selection was made is apparent from a comparison of the average percentages of difficulty of the items for the standardization and experimental groups, which are given in Table 1. The line labeled "E-average" is for the experimental group and is to be compared with the line immediately above which shows the corresponding averages for the standardization group. As was to be expected, the discrepancies are greater for individual items within a set, but even these showed only 4 reversals of rank for a total of 25 items.

6. *Directions to the Subjects.* Every possible pre-

caution was taken to insure adequate comprehension of the tasks by the subjects before the records were taken. Preliminary to reading before the camera, the subject was asked to read and answer six objective questions on a paragraph taken from the same source as the scientific prose used in the experiment. To acquaint the subject with the types of questions he would be expected to answer while the eye-movements were being photographed these questions were formulated so as to represent each type. The paragraph was read silently and the answers to the questions were reported orally to the experimenter. Usually the questions were readily comprehended, but if there was any difficulty the item was carefully explained. Then S was shown a sample set of five questions on a card, exactly as it would appear before the camera, and instructed on how to read it. First, the three numbers at the end of the first line were to be fixated and read slowly; next, question one was to be read and the answer reported orally; and then, instead of going immediately to the next question, the cross at the end of the question was to be fixated until S was told, "Go to the next." Finally, after proceeding through the five questions in this manner, the numbers at the lower left were to be fixated and read slowly. Following this preliminary instruction the subject was seated before the camera, the head-rests were adjusted to hold the head firmly, and the spot of light reflected from the cornea was focused sharply upon the film. At this point an additional practice set of questions was read and answered by the subject while the experimenter observed the

behavior through a peep-hole at the back of the film box. A practice paragraph of scientific prose was read in the same way. When directions were not followed, attention was called to the errors and the sample was repeated. If, however, it was apparent that the subject understood and was following directions, the reading of the scientific prose was photographed. Then always followed the different sets of questions in a systematic order. Just before the reading and answering of each set was photographed, a sample of the corresponding type was exposed in the reading-frame before the subject. This precaution was taken to insure the appropriate "set" for the different types.

The specific directions were as follows: for the prose, "Read the paragraph one time for the content. You will be expected to answer questions on it after you have finished," and for the questions, "Answer each question as quickly as possible but also accurately. If you don't know the answer and are sure you can't figure it out, you may guess, just as you would in a regular examination of this kind."

Consideration of the controls here outlined show that (1) variation in content of material, (2) differences in difficulty from type to type, (3) dissimilarities in the printing set-ups, (4) possible advantages from order of presentation, (5) faulty selection of subjects, and (6) lack of uniformity in directions have been eliminated, or at least have had their influence reduced to a minimum, as possible variables. Under these conditions it appears justifiable to attribute any differences observed in the pattern of eye-movements for the various types of questions to the form in which they are stated.

III

RESULTS[2]

The data obtained by the procedures described above may be presented conveniently under two headings: (1) a quantitative analysis and comparison of the responses to the different types of questions, including a comparison according to levels of difficulty; and (2) a qualitative description of the eye-movement patterns characteristic of each type.

A. QUANTITATIVE ANALYSIS

1. *Comparisons According to Type.* Comparisons are presented for the reading and answering of each type of question, except disarranged sentences,[3] in terms of number of fixations, pause duration, perception time (sum of the pause durations), and frequency of regressive movements. The averages and standard deviations for each of these variables are listed in Table 2. Inspection of this table reveals that the averages for all measures are greater for the questions than for prose, with marked differences among the various types of questions. Before considering the statistical significance of these differences, it may be well to point out some additional facts in the table.

The types are ranked in order of increasing time required to read and answer them as the means in col-

[2]For a complete presentation of raw data and a more comprehensive exposition see thesis on file in the Graduate School, University of Minnesota.
[3]Sufficient data for quantitative comparison are lacking on this type.

TABLE 2
MEANS AND S.D.'S OF READING-MEASURES ON PROSE AND QUESTIONS
$N=66$*

Types	Number of fixations M	Number of fixations S.D.	Pause duration† M	Pause duration† S.D.	Perception time† M	Perception time† S.D.	Number of regressions M	Number of regressions S.D.
Prose	10.44	2.15	12.17	1.27	127.38	31.32	2.11	1.03
True-false	12.05	2.75	13.53	2.36	162.61	50.40	2.84	1.46
Completion	11.12	2.87	16.18	4.19	176.67	70.88	2.75	1.51
Multiple-choice	14.23	3.83	14.06	2.09	198.48	60.80	4.20	1.91
Analogies	16.58	5.63	14.83	1.97	243.08	84.14	5.43	2.93
Wrong-word	18.90	4.70	14.94	1.90	277.60	69.50	6.30	3.08

*In all comparisons N equals 66, except for wrong-word, in which case $N = 50$.
†Time is given in fiftieths of seconds.

umn 6 labeled *Perception time*[4] will show. They are also ranked in order of increasing number of fixations, except that completions should be ahead of true-false questions. This discrepancy in rank for the two constants results from the fact that the pauses in completions are relatively long. Column 4 shows that the pause duration for completions is greater than for any other type. This average pause length should not be interpreted to imply that all the pauses were longer, for the mean is weighted by some extremely long pauses on the blanks at the ends of the questions. This fact, which is indicated by the relatively large standard deviation, will be brought out more emphatically when the plotted records are considered. The column for regressions exhibits the same order as for fixations.

[4]*Perception time* is somewhat less than total reading-time since time for eye-movement is not included. Total reading-time may be approximated by multiplying perception time by 1.06. [Computed from constants reported by Tinker (26).]

The deviations from the regular prose-like progression along the lines of print as shown by the regressions are least in the case of true-false and completion items, the irregularity in sequence of fixations being quite marked for multiple-choice, analogy, and wrong-word questions. Pause durations, except for completions, exhibit the most constant averages and least variability.

Perhaps the coefficients of variation $\left(\frac{\sigma}{m}\right)$ in Table 3 are more interpretive indices of the relative variability of the measures.

TABLE 3

COEFFICIENTS OF VARIATION FOR READING-MEASURES ON PROSE AND QUESTIONS

Types	Fixations	Pause duration	Perception time	Regressions
Prose	.21	.10	.25	.49
True-false	.23	.17	.31	.51
Completion	.26	.26	.40	.55
Multiple-choice	.27	.15	.31	.45
Analogies	.34	.13	.35	.54
Wrong-word	.25	.13	.25	.49

As the table shows, the relative variability for fixations is quite similar for the various types, with the exception of wider variability for analogies. The general spread of measures is least for pause duration, a fact which emphasizes the relatively large variability in the case of completions. Since perception time is the product of number of fixations and pause durations,[5] the greater relative variability might be expected

[5]The means of each of these variables were computed from the raw data and as a result of slight errors in grouping the products do not correspond exactly to the means reported.

to show up in this constant also, and inspection of the table affirms this. Regressions exhibit the widest variation. The significance of these large coefficients will be apparent when the plotted records are considered. To anticipate: they show a range of variation in eye-movements from regular prose-like progression to a series of erratic backward and forward oscillations.

It is pertinent now to inquire about the magnitude and the statistical significance of the differences between the various means in Table 2. Data on these points for each reading-measure are given in Tables 4 to 7, inclusive.

TABLE 4

FIXATIONS: DIFFERENCES BETWEEN MEANS AND $\frac{D}{\sigma_D}$'s*

Figures above the diagonal represent the differences between means; thoes below, the corresponding indices of statistical significance.

R		Prose	T.-F.	Com.	M.-C.	Anal.	W.-W.
(1)	Prose		1.62	.68	3.79	6.14	8.82
(3)	True-false	3.78		.89	2.26	4.59	6.78
(2)	Completion	1.55	2.89		3.11	5.43	7.46
(4)	Multiple-choice	7.54	5.39	8.39		2.32	3.92
(5)	Analogy	8.36	6.76	8.27	4.29		1.40
(6)	Wrong-word	20.00	11.09	11.93	6.97	1.68	

*The formula for correlated measures was used in calculating the σ_D's in this and succeeding tables.

In Table 4 the figures above the diagonal represent the differences between means; those below, the corresponding indices of statistical significance. For example, the difference between the means for true-false and prose is 1.62 fixations per line, and that this difference is statistically significant is indicated by the

ratio 3.78 below the diagonal. The direction of the differences may be conveniently determined by reference to the first column of numbers under *R*, which gives the rank order (smaller number indicates fewer fixations) of each type. Inspection of this table shows that all of the differences, except those between completion and prose, completion and true-false, and wrong-word and analogies, meet the usually accepted criterion[6] for statistical certainty.

Table 5, giving similar data for pause duration, and succeeding tables of differences are to be read and interpreted in the same manner as Table 4.

There are fewer significant differences in pause duration than for fixations. Table 5 shows that between prose and the questions all the differences are significant, in favor of shorter pauses for prose; but for the questions alone six out of ten fail to meet the criterion of statistical significance. This verifies a previous statement concerning the relative constancy of pause duration.

From the point of view of time economy in using objective examination questions, Table 6 is most important. The types, as the table shows, are ranked in order of increasing time required to read and answer them. The greatest difference is between true-false and wrong-word items. The difference between wrong-word and analogies is small and probably not signi-

[6] If $\frac{D}{\sigma_D} = 3$ the chances are 99.9 out of 100 that the true difference is greater than zero (H. E. Garrett, *Statistics in Psychology and Education*, p. 134).

ficant. Similarly, the differences between true-false and completions and between completions and multiple-choice items are of doubtful statistical significance. All other differences, however, are highly significant.

TABLE 5

PAUSE DURATION: DIFFERENCES BETWEEN MEANS AND $\frac{D}{\sigma_D}$'S

Figures above the diagonal represent the differences between means; those below, the corresponding indices of statistical significance.

R	Prose	T.-F.	Com.	M.-C.	Anal.	W.-W.
(1) Prose		1.38	4.01	1.89	2.68	2.78
(2) True-false	5.05		2.58	.49	1.29	1.75
(6) Completion	8.30	5.52		2.12	1.31	1.50
(3) Multiple-choice	8.01	1.85	4.88		.79	.80
(4) Analogy	11.75	4.15	2.54	2.88		.06
(5) Wrong-word	10.90	5.03	2.96	2.75	.21	

TABLE 6

PERCEPTION TIME: DIFFERENCES BETWEEN MEANS AND $\frac{D}{\sigma_D}$'S

Figures above the diagonal represent the differences between means; those below, the corresponding indices of statistical significance.

R	Prose	T.-F.	Com.	M.-C.	Anal.	W.-W.
(1) Prose		35.23	49.10	70.91	115.70	154.40
(2) True-false	5.68		14.31	36.62	81.25	111.20
(3) Completion	5.80	2.00		21.81	66.19	90.40
(4) Multiple-choice	9.14	5.38	2.67		44.47	66.00
(5) Analogy	10.92	7.95	5.61	5.71		19.20
(6) Wrong-word	16.19	11.89	8.17	9.52	1.64	

These rankings agree with those reported by Ruch (23) for recall, multiple-choice, and true-false questions, except that "recall," which corresponds to com-

pletions in this study, required the longest time. A part of this discrepancy is doubtless explainable by the fact that subjects in the studies reported by Ruch were required to write out the answers, whereas here they reported verbally. The results obtained by Eurich (13), however, indicate that for some kinds of material the difference between true-false and completion questions is slight, a finding confirmed by the present study.

TABLE 7

REGRESSIONS: DIFFERENCES BETWEEN MEANS AND $\frac{D}{\sigma_D}$'s

Figures above the diagonal represent the differences between means; those below, the corresponding indices of statistical significance.

R	Prose	T.-F.	Com.	M.-C.	Anal.	W.-W.
(1) Prose		.75	.64	2.09	3.32	4.42
(3) True-false	3.36		.07	1.41	2.63	3.43
(2) Completion	2.91	.35		1.45	2.66	3.40
(4) Multiple-choice	8.10	6.40	7.44		1.21	1.83
(5) Analogy	8.74	7.73	8.72	3.65		.42
(6) Wrong-word	11.00	8.55	8.34	4.48	.96	

Table 7 shows marked differences in the regularity in sequence of fixations for all question comparisons, except between true-false and completions and between wrong-word and analogy. This table also reveals the single instance in which a measure between prose and the questions is not significant—the difference between prose and completions is only slightly less than that required for statistical significance.

In order to make the relative magnitudes of these differences more comprehensible, and by way of sum-

mary, all the measures have been converted into percentages and brought together in Table 8.

TABLE 8
PERCENTAGE WHICH EACH AVERAGE IS OF THE CORRESPONDING WRONG-WORD AVERAGE

Types	Number of fixations	Pause duration	Perception time	Number of regressions
Prose	55	82	46	34
True-false	64	90	59	45
Completion	59	108	64	44
Multiple-choice	75	94	72	67
Analogy	88	99	88	86
Wrong-word	*100*	*100*	*100*	100

The interpretation of the table needs no elaboration. Column 4, perception time, shows the fractional amounts of time required to answer each type in terms of wrong-word items. This corresponds closely to the number both of fixations and of regressions which occur in reading each type. Completions in each instance constitute the exception to perfect agreement in rank. The reason is apparent in column 3, which indicates the relatively long average pause duration *for this type*. Completions, however, as has already been pointed out, are read like prose, with the exception of relatively long end pauses which contribute disproportionately to the means for pause duration and also to average perception time. *Incidentally*, this column again emphasizes the high degree of constancy in pause duration from type to type.

2. *Correlations*. An attempt has been made to ascertain the relative degrees of interrelationship among

the variables by a study of the intercorrelations. These indices for the four measures of each type and prose are presented in Tables 9 to 12.

TABLE 9
FIXATIONS: PEARSON r'S AND P.E.'S

Types	T.-F.	Com.	M.-C.	Anal.	W.-W.
Prose	.34±.07	.15±.09	.16±.08	.06±.08	.22±.09
True-false		.61±.05	.46±.07	.30±.08	.42±.08
Completion			.63±.05	.37±.07	.41±.08
Multiple-choice				.63±.05	.58±.06
Analogy					.39±.08

The absolute values of the coefficients in Table 10 and those included in subsequent tables are unimportant; they are undoubtedly attenuated by somewhat low reliabilities—a point to be considered later. The relative differences, however, appear significant. In general, the correlations between each type and prose are very low; and, in contrast with these, the intercorrelations among the questions are fairly marked. This general difference probably reflects a modification in attitude which remains somewhat similar for the various questions but differs from that characteristic of prose reading. The single exception is the correlation between prose and true-false items, which is slightly higher than the true-false–with–analogies coefficient.

TABLE 10
PAUSE DURATION: PEARSON r'S AND P.E.'S

Types	T.-F.	Com.	M.-C.	Anal.	W.-W.
Prose	.39±.07	.35±.07	.44±.07	.42±.07	.43±.08
True-false		.45±.07	.54±.06	.36±.07	.42±.08
Completion			.54±.06	.26±.08	.60±.06
Multiple-choice				.41±.07	.49±.07
Analogy					.41±.08

The difference between prose and the questions does not hold for pause duration, as is apparent in Table 10. This, however, was to be expected from previous observations on the relative constancy of pause duration.

Table 11 shows the same general trend as observed for fixations, though the difference between prose-question and inter-question coefficients is less marked. This, of course, was to be expected since pause duration enters into perception time.

TABLE 11

PERCEPTION TIME: PEARSON r'S AND P.E.'S

Types	T.-F.	Com.	M.-C.	Anal.	W.-W.
Prose	.33±.07	.29±.08	.18±.08	.15±.08	.28±.09
True-false		.61±.05	.53±.06	.31±.08	.45±.07
Completion			.50±.06	.26±.08	.42±.08
Multiple-choice				.56±.06	.72±.04
Analogy					.46±.07

TABLE 12

REGRESSIONS: PEARSON r'S AND P.E.'S

Types	T.-F.	Com.	M.-C.	Anal.	W.-W.
Prose	.10±.08	.13±.08	.14±.08	.08±.08	.35±.08
True-false		.59±.05	.38±.07	.37±.07	.39±.08
Completion			.59±.05	.55±.06	.38±.08
Multiple-choice				.45±.07	.40±.08
Analogy					.49±.07

Inspection of Table 12 shows the same general trend for regressions that was observed for fixations and perception time. The prose-question correlations, with the exception of prose *vs.* wrong-word, are near zero; while, relatively, the inter-question coefficients are fairly high.

3. *Reliabilities.* Determination of the reliabilities of the eye-movement measures obtained for the various

types of questions is a difficult problem. For the prose the sample of reading is fairly adequate; but each set of questions includes only five items and these differ in difficulty. However, all the comparisons have been made with group averages, and, with respect to the correlations, only relative differences have been utilized. Nevertheless, an estimate of reliability has been attempted and the results are presented in Table 13.

TABLE 13

RELIABILITIES: PROSE—ODDS *vs.* EVENS; QUESTIONS—SECONDS *vs.* FOURTH QUESTION IN EACH SET

Types	Fixations		Perception Time*		Regressions	
	Raw *r*	S-B *r*	Raw *r*	S-B *r*	Raw *r*	S-B *r*
Prose	.77	.87	.79	.88	.64	.78
True-false	.28	.66	.42	.78	.47	.81
Completion	.31	.70	.33	.71	.22	.58
Multiple-choice	.42	.78	.42	.78	.47	.81
Analogy	.33	.72	.42	.78	.20	.55
Wrong-word	.46	.81	.27	.65	.65	.90

*Pause duration has not been included since it is combined with number of fixations in perception time.

The first column under each heading gives the raw coefficient for each type of material. For the prose, these indicate the reliability of half the material (odd *vs.* even lines) and for the questions, of one-fifth (second *vs.* fourth items). The second columns under the corresponding headings show the correlations, estimated by the Spearman-Brown prophecy formula, for the amount of material actually used.

4. *Comparison According to Levels of Difficulty.* Since there are five measured levels of difficulty represented in each set of questions it was possible to deter-

mine to what extent these differences conditioned variations in the eye-movement measures. It might have been expected that the harder the question the greater would be the effort required to read and answer it; and this should have been reflected in increased number of fixations, regressions, and perception time. However, except for very easy and harder items, the records show that the mean differences between successive levels are slight. For each type of question, fluctuation from one level to another in the eye-movement measures occur, but the direction of change is inconsistent and probably the result of chance errors of sampling; it does not parallel the increments of difficulty, with this exception: In every case, except one, the easiest item requires fewer fixations than any harder question, and always the hardest item exceeds the easiest in average number of fixations. Moreover, the differences between the easiest and hardest items are statistically significant for all types except analogies.

Incidentally, this discovered lack of correspondence between difficulty and the eye-movement measures supports the validity of the differences found for the various types of questions by showing that the preciseness of the control for difficulty was more than sufficient. The differences between significantly different types exceed the range of variation according to levels of difficulty.

5. *Summary.* The various types of questions have been compared with each other and with prose in terms of averages for number of fixations, pause duration, perception time, and frequency of regressive move-

ments. Between the questions as a whole and prose the differences are marked and statistically significant in every instance except one—the number of fixations and regressions required to read and answer completions is only slightly greater than for prose. In terms of these measures, true-false items differ little from completions, and wrong-word questions are very similar to analogies. Pause duration, which is the least variable of all measures, exhibits little difference from type to type, except that on the average it is unusually long for completions. One notes especially the marked differences between the questions and prose. For perception time, which is most important from the point of view of time economy in the use of objective questions, the rank order is as follows: true-false, completions, multiple-choice, analogy, and wrong-word items. The differences are statistically significant in every case except true-false *vs.* completions, multiple-choice *vs.* completions, and analogy *vs.* wrong-word. All the types differ significantly from prose. The mean comparisons for regressions which are indices of the irregularities in the sequence of eye-movements show that completions and true-false items are read much like prose while the other types are read much more analytically, particularly analogies and wrong-word questions.

An analysis of the intercorrelation coefficients reveals a quite consistent general trend. Practically all the correlations between prose and the various types, except for pause duration, are near zero; in contrast with this, the question intercorrelations are, relatively, fairly high. This is interpreted as indicative of a

change in attitude which is modified in going from prose to questions but remains quite similar from one type of question to another.

Comparisons of questions varying in degree of difficulty reveal no differences in average number of fixations, perception time, and average frequency of regressions, which may be considered to parallel successive increments of difficulty, except between the easiest and harder items.

B. Qualitative Description of How Objective Questions Are Read

The qualitative description of how the different types of questions are read is chiefly in terms of pattern or sequence of eye-movements, but number and durations of pauses have also been considered. The plotted records, illustrating the various procedures employed in reading the different forms of questions, constitute the basis for description.

Perhaps the most striking fact revealed by the plotted records is the breadth of variety. Variation in the patterns of eye-movements both for different types and for different questions of the same form ranges from the smooth sequence of pauses typical of easy prose to an erratic series of oscillatory backward and forward movements, which defies classification as "reading." Nevertheless, a kind of classification, somewhat arbitrary, to be sure, is possible. In fact, when the various procedures are grouped into categories displaying rather marked similarities, consider-

able consistency is apparent. The nature of this consistency and the basis for the classification can be illustrated best by presenting the classification itself.

The classified material includes 170 plotted records, 30 of each type except disarranged sentences, for which only 20 records were plotted. These records, consisting of an easy, average, and most difficult item from each type, represent the performances of only 10 subjects. They are probably typical of the entire group, however, because the criteria for selection were clearness and distinct separation of dots on the film. Since these factors are determined by the experimenter and apparatus (the light might vary in intensity, the speed of the passage of film might change slightly, etc.), the selection of subjects would be perfectly random.

In classifying the modes of procedure in reading and answering the questions, records were considered[7] "analytical," "erratic," etc., relative to the corresponding typical prose performance for each individual. Samples of the eye-movement records in reading prose for the 10 subjects whose question records are plotted in Figures 2 to 7 are presented in Figure 1. The prose and question records for a given subject may be identified by the number preceding each item in the column headed "*S*."

Inspection of Figure 1, showing the sample prose records, indicates a fairly wide extent of variability. Since the line for each subject was chosen to represent

[7]Classification is based on subjective estimates of the author; for independent verification see complete original data in appendix to thesis on file in Graduate School, University of Minnesota.

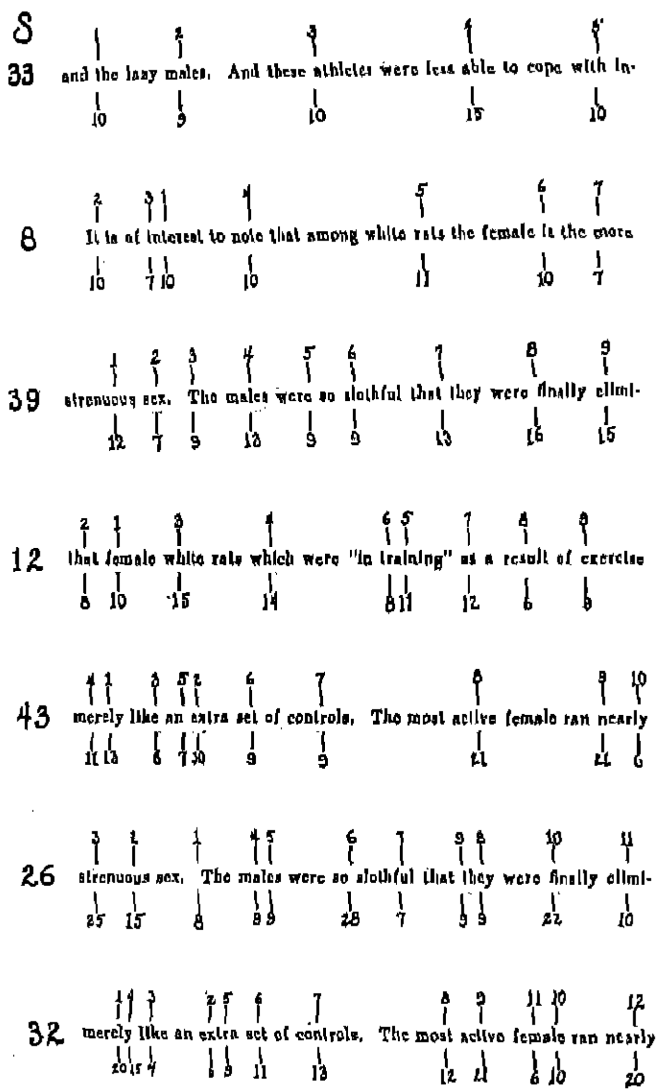


FIGURE 1

Representative Prose for Subjects Whose Question Records Are Plotted in Figures 2 to 7

his average performance as nearly as possible, this variation may be considered to characterize the group as a whole. (Only 7 records are reproduced, but since Number 22 is similar to 8, 50 similar to 39, and 30 similar to 12, one record of each pair may be taken as representative of the other.) The series of numbers above each line show the order of fixations and the corresponding figures below the line give the durations of each pause in fiftieths of seconds. For example, in the record of Subject 33, showing a rapid sequence of eye-movements, the first pause at the beginning of the line lasted 10/50 of a second, the second 9/50, etc.

The plotted points represent exact points of reflection of light from the "bright spot" on the cornea, but they should not be interpreted to imply that the corresponding points of visual fixation are so sharply centered. Dodge (11) and others (8, 25) have shown that it would be more exact to speak of "fixation field" rather than point. This field extends over an area of 3 or 4 letter spaces (25), exclusive of "marginal impressions" (11, 15) from eccentric stimulation. Thus the sensory effects from fixation on any point or even near a word are practically equivalent.

With this orientation we consider now the description of the various types of questions.

1. *Wrong-Word.* Representative procedures in reading wrong-word statements are illustrated in Figure 2. It is apparent from inspection of the records plotted in this plate that the reading of wrong-word questions is characterized by at least six varieties

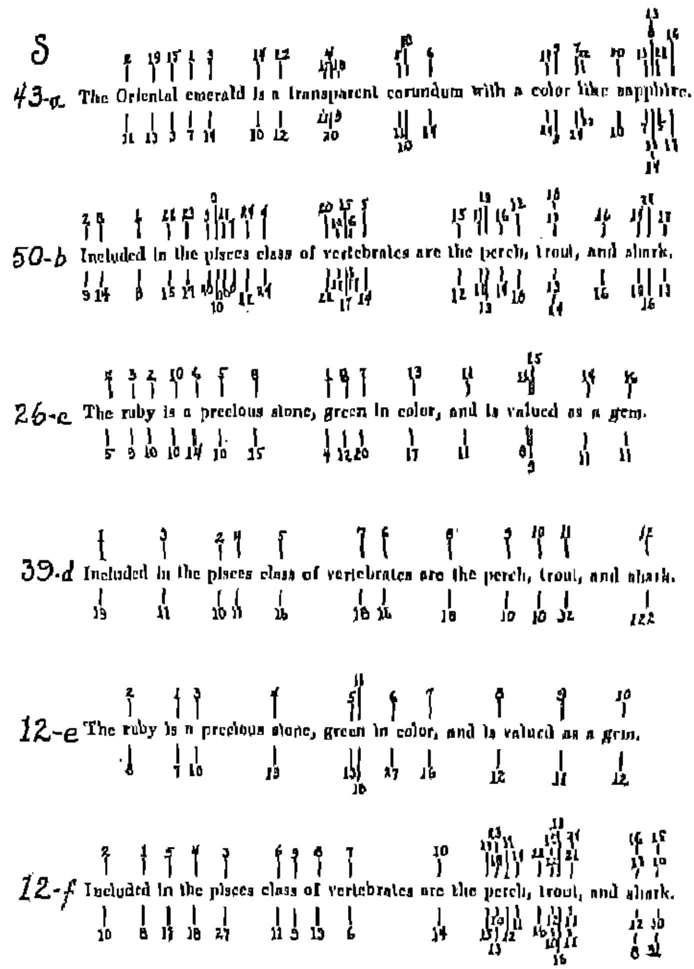


FIGURE 2

REPRESENTATIVE PROCEDURES IN READING WRONG-WORD STATEMENTS

(*a*) regular progression first, then analytical, 13;* (*b*) analytical, 6; (*c*) "plodding" analytical progression, 3; (*d*) careful regular progression, 3; (*e*) regular progression with "recheck" on answer, 3; (*f*) analytical progression with concentration on significant words, 2.

*The numbers indicate the frequency with which each variety occurred in a total of 30.

of eye-movement patterns. Typically, in 43 per cent of the cases, the questions are read through first with a regular prose-like sequence of eye-movements, and then follows an analytical or irregular oscillating sequence. The first line in the plate, 43-*a*, illustrates this mode of performance. Line 50-*b* shows a type, representative of 20 per cent of the records, where there is no preliminary "regular" reading but only an irregular sequence from beginning to end—the process is analytical throughout. A third variety, line 26-*c*, which occurs less frequently, consists of a "plodding" analytical progression directly along the line of print; there are regressions, but they are much shorter than those in the two preceding lines. A fourth type, line 39-*d*, occurring with the same frequency as 26-*c* (10 per cent), illustrates what has been called "careful regular progression"; there are no more regressions than in prose reading, and it is distinguished from the latter only by more and longer fixations. In a fifth type, 12-*e*, also occurring with a frequency of 10 per cent, the questions are read through like line 43-*a* in regular prose fashion and then there is a refixation on the word reported as the answer. Finally, a variety was noted distinguished by analytical progression through the statement followed by concentration on significant words, as illustrated by 12-*f*.

Since the varieties of procedure overlap the degrees of difficulty of the items, all three levels of difficulty are found frequently in each category. Therefore, this factor has been neglected and the questions have been classified only according to type and variety of eye-movement pattern.

2. *Analogies.* The variation in pattern of procedure is even greater for analogies than for wrong-word statements. This, however, was to be expected because of the more complex structure of the questions. Wrong-word items, being simply prose statements, are

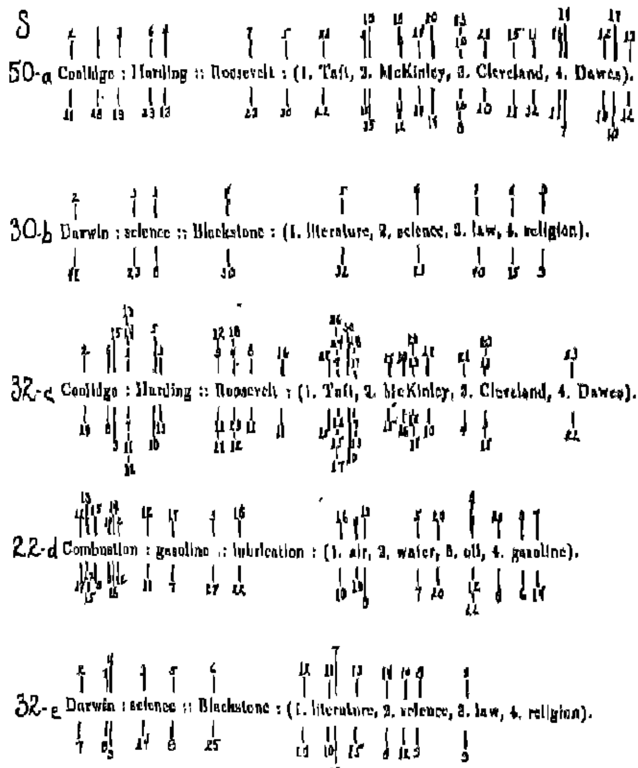


FIGURE 3a

Representative Procedures in Reading Analogies

(*a*) analytical progression, relation first and then alternatives, 6; (*b*) regular careful progression, 6; (*c*) analytical or erratic without direct progression, 6; (*d*) scan and then erratic, 4; (*e*) analytical progression to alternatives and then analytical with incomplete reading, 2.

read differently because of the changed attitude evoked by the directions to look for a specific item of content, namely, the word which makes the statement wrong. The analogy consists really of two parts: a given relationship, more or less complex in itself, and a corre-

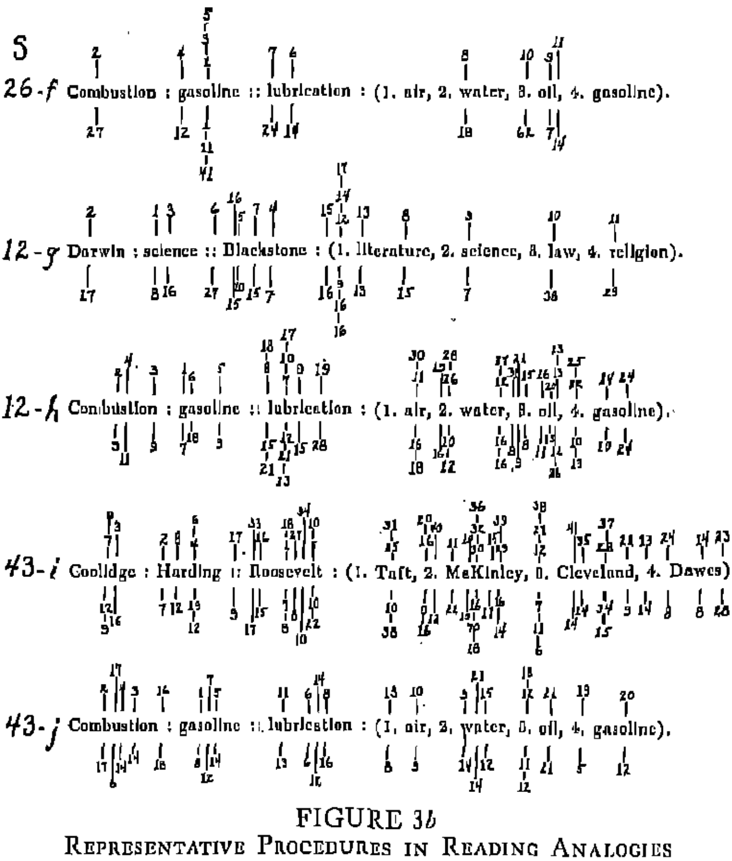


FIGURE 3*b*

REPRESENTATIVE PROCEDURES IN READING ANALOGIES

(*f*) analytical progression to alternatives, then "spot" and concentration, 2; (*g*) analytical progression to alternatives, then careful progression and concentration, 1; (*h*) analytical to alternatives, then careful progression, and finally erratic, 1; (*i*) analytical progression and through alternatives, then erratic, 1; (*j*) erratic, 1.

late which is to be deduced by choosing the fitting word from four alternatives.

Ten different modes of reacting to this type of question were noted, examples of each variety being exhibited in Figures 3*a* and 3*b*. The procedures are typically analytical. Twenty per cent of the analogies, 50-*a*, show analytical progression through the first relationship and then an analytical irregular series of fixations distributed over the alternatives. Another 20 per cent were read in regular, careful, prose-like progression, 30-*b*. A third type, 32-*c*, occurring with equal frequency, illustrates an analytical series of forward and backward oscillations without direct progression. Item 22-*d*, however, shows a rapid "scanning" progression followed by erratic oscillatory movements. A fifth type, 32-*e*, exhibited in 6 2/3 per cent of the records, differs from other analytical varieties only in that the analytical progression is terminated without fixations on all the alternatives. Another type, with the same frequency, 26-*f*, also illustrates analytical progression to the alternatives, then a "spotting" procedure, in which perception was evidently mediated through eccentric stimulation, and, finally, concentration upon the choice of alternatives. Four other analytical types, each with a frequency of 3 1/3 per cent show, respectively: analytical progression to the alternatives succeeded by careful progression and concentration (12-*g*); analytical progression to the alternatives and then careful progression followed by erratic oscillations (12-*h*); analytical progression to and through the alternatives and then erratic eye-

movements (43-*i*); and, finally, a type characterized by an erratic sequence of eye-movements throughout.

3. *Multiple-Choice.* Like the analogies, the multiple-choice items are constructed with two parts. The statement preceding the alternatives in the multiple-choice question, however, is much less complex than the first part in the analogy, and for this reason some differences in eye-movement patterns would be expected. Both types are usually read analytically, but there is a difference in degree, the multiple-choice items showing the more regular sequence of eye-movements. The nine varieties of procedure which are displayed indicate a similar degree of variety.

Most frequently, 27 per cent of the records, the multiple-choice questions are read with analytical progression to and through the alternatives, but with more concentration on the alternatives. Item 30-*a*, Figure 4-*a*, illustrates this type. In another kind, 22-*b*, occurring with a frequency of 23 per cent, there is first regular prose-like progression to the alternatives and then analysis. Question 26-*c* shows a feature noted in the analogies but more characteristic of multiple-choice items, namely incomplete reading of the alternatives. In this item regular progression to the alternatives terminates with a concentration on the answer reported, without fixation on the remaining alternatives. Three other varieties of multiple-choice procedures, differing in other respects, exhibit this same phenomenon. In 50-*d* analytical progression to the alternatives is followed by concentration on the first alternative which

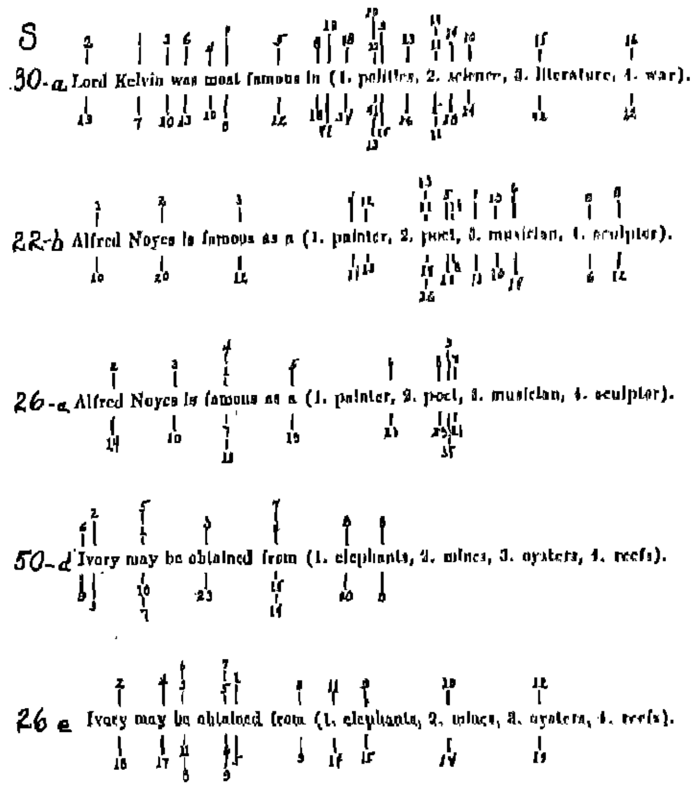


FIGURE 4*a*

REPRESENTATIVE PROCEDURES IN READING MULTIPLE-CHOICE QUESTIONS

(*a*) analytical progression to and through alternatives, 8; (*b*) regular progression to alternatives, then analytical, 7; (*c*) regular progression with incomplete reading and concentration on choice of alternatives, 4; (*d*) analytical progression with incomplete reading and concentration on choice, 4; (*e*) analytical progression with incomplete reading of alternatives, 1.

was reported as the answer without further fixations. Line 26-*b* differs from 50-*d* only in the absence of concentration on the alternative answered. Line 32-*f*, Figure 4*b*, illustrates a similar partial reading of the

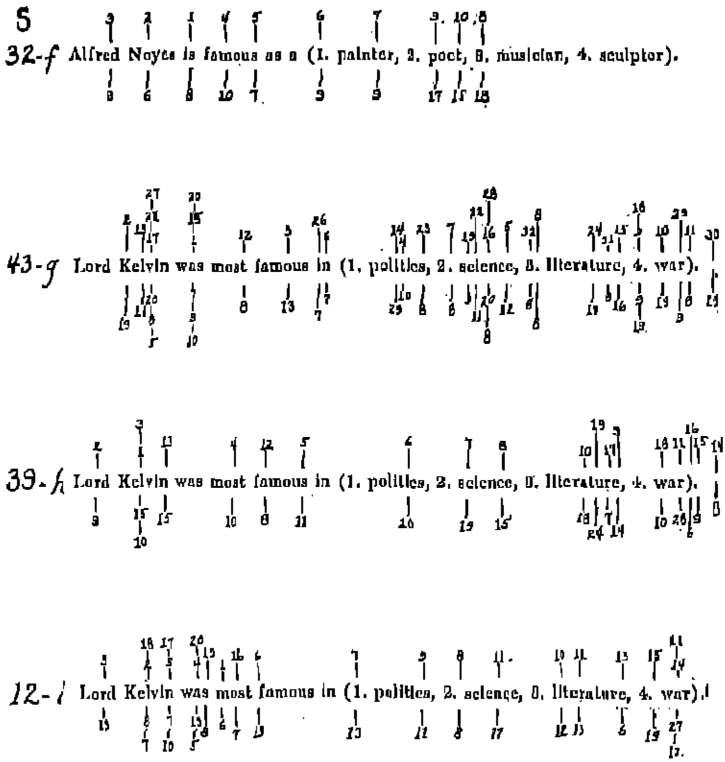


FIGURE 4*b*

REPRESENTATIVE PROCEDURES IN READING MULTIPLE-CHOICE QUESTIONS

(*f*) regular progression with incomplete reading of alternatives, 2; (*g*) erratic, 2; (*h*) regular progression to and through alternatives, then erratic, 1; (*i*) analytical progression to and through alternatives, then erratic, 1.

alternatives without concentration, but the first part of the question is read more like prose. Sometimes, as shown by line 43-*g*, the sequence of eye-movements is erratic from beginning to end. Item 39-*h* illustrates regular progression to and through the alternatives and

then erratic to and from oscillations. And, finally, as in 12-*i*, the erratic sequences of eye-movements are preceded by analytical progression to and through the alternatives.

4. *True-False.* One should expect true-false items

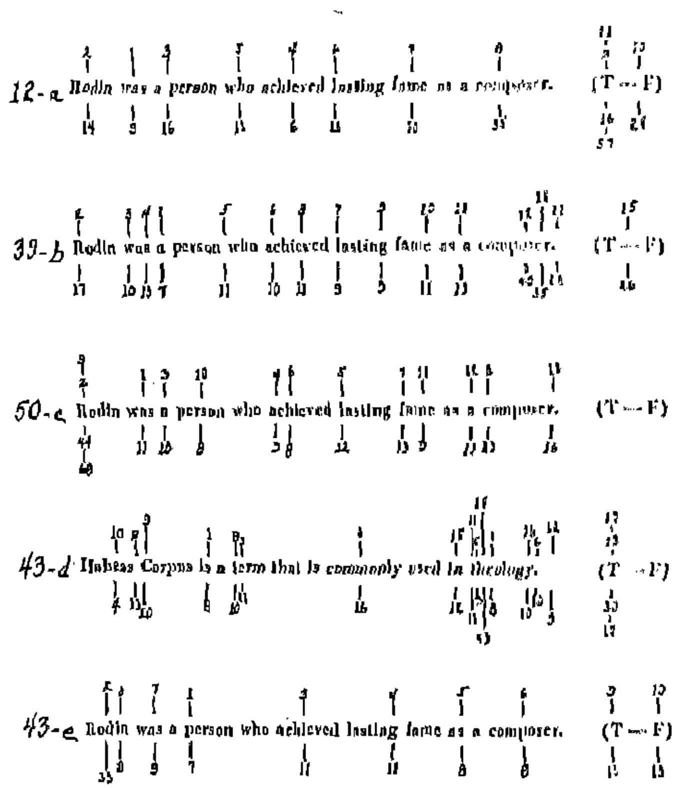


FIGURE 5

REPRESENTATIVE PROCEDURES IN READING TRUE-FALSE STATEMENTS

(*a*) regular progression with pause on "T-F," 17; (*b*) analytical progression with pause on "T-F," 6; (*c*) analytical, 3; (*d*) regular progression, then analytical with pause on "T-F," 2; (*e*) regular progression with "recheck" and pause on "T-F," 2.

to be read less analytically than the previously considered types since attention is not necessarily directed to specific items of content but more to comprehending the statements as wholes, as is true for prose. The typical reading of this type (57 per cent of the cases) is shown by 12-*a* in Figure 5, where a regular prose-like sequence of eye-movements is followed by a rather long pause on the "T-F" abbreviation. Another 20 per cent, however, are read analytically with a similar pause at the end, 39-*b*. Another 10 per cent differ from line 39-*b* only in the absence of the end pause, 50-*c*. In 6 2/3 per cent of the instances the prose-like progression during the first reading changes to the analytical type, 43-*d*. And, finally, as in 43-*e*, occurring with equal frequency, regular progression is followed by a short refixation on a significant word and then a pause at the end.

5. *One-Word Completions.* Simple completion questions, like true-false statements, should be expected to be read much like prose since the solutions to the problems cannot be found in further analysis after the question is once comprehended. The fact that even less variation of procedure occurs with this type than for true-false statements supports this supposition. Line 30-*a*, Figure 6, displays the typical procedure for completions. This regular prose-like progression of eye-movements with a relatively long pause on the blank at the end represents 63 per cent of the readings. Sometimes (13 per cent), however, as shown by 22-*b*, the regular progression stops without an end pause. Equally frequent is the type, 39-*c*, exhibiting analytical

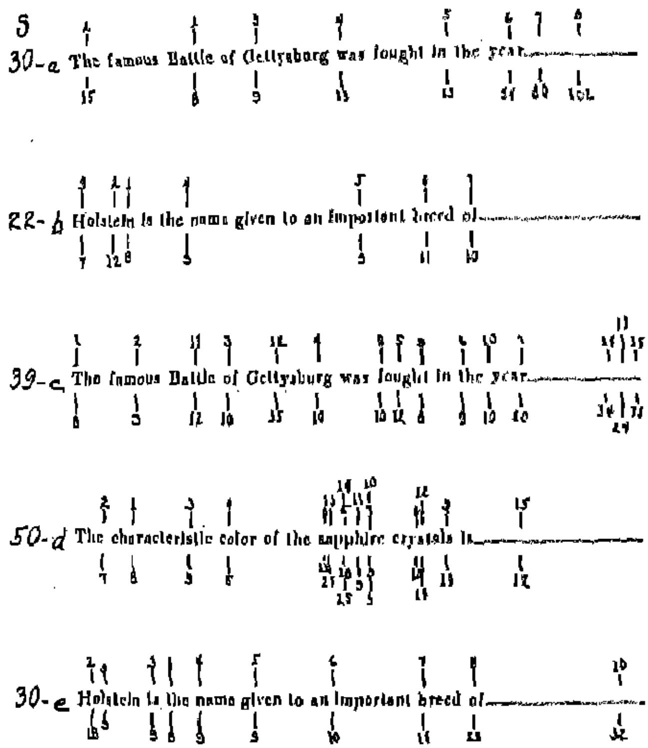


FIGURE 6

REPRESENTATIVE PROCEDURES IN READING COMPLETION QUESTIONS

(*a*) regular progression with pause on blank, 19; (*b*) regular progression, 4; (*c*) analytical progression with pause on blank, 4; (*d*) regular progression with pause on significant word, 2; (*e*) regular progression with "recheck" and pause on blank, 1.

progression with an end pause on the blank. In fewer instances (6 2/3 per cent) one finds regular progression with a concentration of pauses on a significant word, 50-*d*. Finally, as was noted for true-false statements, in a few instances (3 1/3 per cent), 30-*e*, a regular

sequence of eye-movements followed by a short refixation and an end pause occurs.

6. *Disarranged Sentences.* It has been observed that an analytical or even erratic sequence of eye-movements is much more characteristic of wrong-word, analogy, and multiple-choice items than it is of true-false and completion statements. The task set by the disarranged sentence, where the attention is directed to specific details in the item, also should be expected to evoke an analytical procedure. The

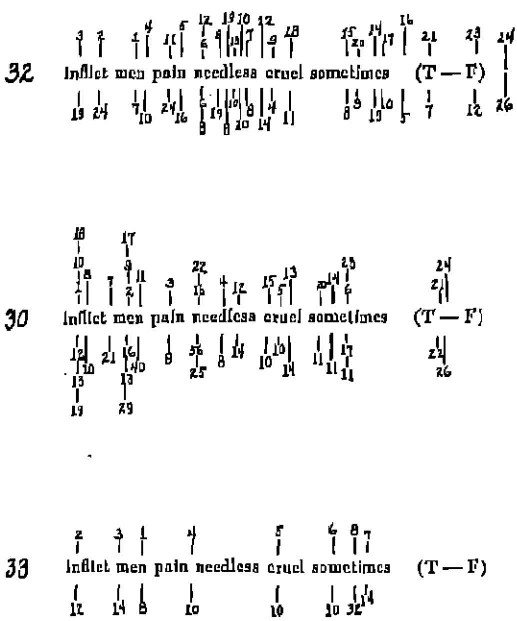


FIGURE 7

REPRESENTATIVE PROCEDURES IN READING DISARRANGED SENTENCES

(*a*) analytical, 12; (*b*) regular progression first, then analytical, 5; (*c*) careful progression, 3.

plotted records presented in Figure 7 show that an analytical procedure characterizes 85 per cent of the records. The most frequently occurring type (60 per cent) is illustrated by item 32-*a*, which shows analytical procedure throughout. In 25 per cent of the records the question is first "surveyed" with a regular sequence of eye-movements succeeded by an erratic sequence, 30-*b*. Finally, a third variety exhibits a "careful" but regular succession of fixations, 33-*c*.

7. *Summary and Discussion.* The eye-movement patterns exhibited in reading and answering objective questions differ most from those characteristic of prose in the more analytical character of the former; for the questions the sequence of fixations is more irregular and pauses are both more frequent and longer. A second striking general difference is the variety of procedures exhibited in reading the different types of questions. The patterns range from a regular prose-like progression to an irregular series of oscillations suggestive of the reading of the isolated formulae in Tinker's study, which he characterized as being "read much as a design is studied" (21). Between these extremes are the typical modes for the various types.

A prose-like progression followed by an analytical series of saccadic shifts most frequently characterizes the reading of wrong-word items, although some are read analytically throughout. Analogies, distinguished by the greater variety of procedure, typically show irregular, analytical sequence. Multiple-choice questions, close rivals for variety, also are usually read analytically, but regular progression frequently occurs,

particularly on the first part. True-false and completion questions, most similar to prose, exhibit the least variability. More than half the items are read like prose, although some analytical reading occurs. This similarity of completions and true-false statements to prose, which is in contrast to the marked differences between prose and the other types in the matter of regressions, is consistent with the differences revealed by the quantitative comparisons. The reading of disarranged sentences is almost wholly analytical.

Of the more specific features observed, the long pauses and partial readings are especially interesting. These long pauses occur at the ends of true-false items, and more frequently on the blanks at the end of completions. Fixations lasting as long as 9 seconds, to mention an extreme, account for the comparatively high average pause duration and relatively greater dispersion previously noted for completions. The pauses in these types indicate that once the question is comprehended, further perception is unnecessary. In the wrong-word, analogy, and multiple-choice questions, however, where in every instance new fixations occur continuously until the solution is reached, it appears that the central processes of assimilation and association require constant perceptual supplementation. This difference may be significant for the consideration of any possible specific validity possessed by the different types.

Incidentally, the relatively limited perceptual behavior required to answer true-false statements may explain the failures of Stumpf (28) and others to dis-

cover any difference in response to these items whether presented orally or visually.

Partial or incomplete readings, most typical of multiple-choice items, also occur with analogies. Since these two types are alike in the form of the alternatives, the reason for the more frequent occurrence of this phenomenon with multiple-choice questions must be due to differences in the first part of the statements. The greater simplicity of multiple-choice statement probably initiates a more definite "set" for the choice, and consequently the correct alternative is more easily recognized. These incomplete readings, which occur whether the correct alternative is in the first, second, or third position, suggest a nice adjustment of eye-movements to the demands of a particular situation. Eye-movements, it appears, are made in response to the needs of the central processes of assimilation and association, the responses to the external stimuli being secondary.

It should be noted, however, that failure to fixate upon a word, as Dodge (11) and Hamilton (15) have shown, does not necessarily imply that it has not been perceived. "Marginal impressions" from eccentric stimulation may have been sufficient for recognition of non-fixated words. But marginal stimuli normally initiate reflex movements of the eye in their direction (15); these, however, are inhibited in the case of the partial readings. Therefore, whether the non-fixated words are regarded as neglected or eccentrically perceived, we have the apparent demonstration of central dominance over the peripheral processes.

The long pauses, as has already been indicated, also illustrate an effective adjustment of the motor to the central processes. Once the true-false or completion statement is apprehended, the mental associations take place without further perception, and perhaps easier, since the pauses presumably involve less distraction than the fixation of new material.[8]

These special instances, however, are not the only data to support an hypothesis of central dominance. The variations in pattern of eye-movements corresponding to the different types of questions indicate adjustments, which logically, at least, appear suited to each particular situation. Even the most erratic oscillations could hardly be called ineffective for the situations where they occur since they are accompanied more often than not with correct responses to the questions. (It will be recalled that the percentage of incorrect answers averaged about 40, and that the average perception time was only slightly greater for incorrectly answered items. Moreover, some of the items scored wrong were undoubtedly right according to the misinformation of the subjects.) Certainly in these instances it would be misleading to interpret an erratic series of eye-movements as indicative of "confused mental processes." In reading easy prose erratic eye-movements may be symptomatic of mental confusion, but in answering objective questions, paraphrasing, reading formulae, and even in reading Latin,

[8]This interpretation was suggested by Buswell's study (4).

where the grammar is highly formalized, they may represent an adequate adjustment to the situations.

Finally, the general trend of results from previous investigations of special types of reading is in harmony with the view here advanced. In some of these studies the changes in purpose or content of reading are so slight that they require only slight modification of the eye-movement pattern which characterizes the reading of easy prose; in others, the changes become so radically different that prose habits are completely abandoned. But in every case a rapid and more or less effective adjustment to each situation is typical. The fact that habits acquired from constant reading of prose do not interfere when the situation demands a different pattern of eye-movements suggests that the oculomotor processes are subordinated and flexible in their responses to the needs of the central processes of assimilation and association.

IV

SUMMARY

This study attempts to describe in both quantitative and qualitative terms how objective examination questions differing in type and by measured amounts of difficulty are read, as they are answered. The eye-movements, which constitute the units of description, were photographed with a modification of the Dodge apparatus. The material read included a paragraph of scientific prose and the following types of objective examination questions: wrong-word, analogy, multiple-choice, true-false, completion, and disarranged sentences. Photographic records, yielding measures of frequency of fixations, duration of pauses, and sequence of eye-movements, were obtained for 66 college sophomores.

The controls were such that (1) variation in content of material, (2) differences in difficulty from one type to another, (3) dissimilarities in the printing set-up, (4) possible advantages from order of presentation, (5) selection of subjects, and (6) lack of uniformity in directions have been eliminated or at least have had their influence as possible variables reduced to a minimum. Under these conditions it appears justifiable to attribute any differences observed in the pattern of eye-movements for the various types of questions, and in comparison with prose, to the effect of the form of the questions. And within each set of questions where difficulty is the experimental variable, other factors

being controlled, possible differences may be concluded to have resulted from the influence of this factor.

A quantitative analysis of eye-movement measures justifies the following statements:

1. In every instance the prose averages for number of fixations, duration of pauses, perception time, and frequency of regressions are exceeded by the corresponding means for the questions. The relative magnitudes of the differences vary, but they are statistically significant in every case except the prose-completion comparisons for fixations and regressions. This apparent discrepancy is consistent with the special sequence of eye-movements characteristic of this type of question.

2. The types of questions are ranked in order of increasing perception time as follows: true-false, completion, multiple-choice, analogy, and wrong-word. The differences between ranks are all statistically significant, excepting the true-false *vs.* completion, completion *vs.* multiple-choice, and analogy *vs.* wrong-word comparisons.

3. The various types take the same ranks for fixations, except that completions require slightly fewer fixations than true-false statements. This pair and the analogy and wrong-word pair are the only comparisons which fail to meet the criterion of statistical significance.

4. Average pause duration varies only slightly from one type to another. The rank orders are the same as those for perception time, except that comple-

tions rise from second to fifth place, but only 4 out of the 10 differences between types are statistically significant.

5. The ranks according to average number of regressions for each type parallel the rank order for fixations, and the same differences, true-false *vs.* completions and analogy *vs.* wrong-word, are statistically not significant. The true-false and completion questions differ markedly less from prose in number of regressions than do multiple-choice, analogy, and wrong-word items.

6. The relative variability of the eye-movement measures, as expressed by the coefficients of variation, is fairly constant from one type to another for fixations, perception time, and regressions. In pause duration, however, the relative variability is significantly greater for completions than for other types. In general, regressions exhibited the largest extent of variation and pause duration the smallest.

7. The intercorrelations reveal a quite consistent trend. Practically all the correlations between prose and the various types, except for pause duration, are near zero; in contrast with this, the question intercorrelations are, relatively, fairly high. This general difference probably reflects a modification in attitude which remains somewhat similar for the various questions but differs from that characteristic of prose reading.

8. In general the relationship is slight between levels of difficulty (5 levels are represented) of the

questions and the eye-movement measures. Very easy items require fewer fixations and regressions and shorter perception time than harder items, but only between the easiest and the hardest is there a consistent significant difference.

A qualitative analysis of 170 plotted records reveals such variety in reading-procedures that any brief summary of the results must be limited in preciseness. The general characterizations listed below mask significant details; they indicate only the most frequently noted features.

1. Wrong-word questions are usually read with careful prose-like progression followed by an analytical series of eye-movements. Sometimes, however, they are read analytically throughout.

2. Analogies, for which the greatest variety of procedures occur, are typified by irregular, analytical sequences of fixations.

3. Multiple-choice items, close rivals of analogies for variety of procedure, are most frequently read analytically, but regular prose-like progression also occurs, particularly on the first part. Partial or incomplete reading of the alternatives is frequently exhibited with this type and occasionally with analogies.

4. True-false and completion items show the greatest similarity to prose and exhibit the least variability in sequence of fixations. Analytical reading occurs, but most frequently, except for the endings, they are read like prose. A feature specifically characteristic of completions is a long pause on the blanks at the ends of the statements.

5. Disarranged sentences are generally read analytically throughout.

The two most general differentiae distinguishing the reading of objective questions from prose reading are the more analytical sequence of eye-movements and the striking variety of procedures employed in reading the questions. In each instance, however, the mode adopted for a particular type appears suited to the situation. Notwithstanding constant practice with prose, this variety of apparently effective procedure and such special features as the long pauses on completions and the partial reading of multiple-choice and analogy questions indicate that the oculomotor processes are subordinated and flexible in their responses to the central processes of assimilation and association.

References

1. Abernathy, E. M. Photographic records of eye movements in studying spelling. *J. Educ. Psychol.*, 1929, 20, 695-701.
2. Buswell, G. F. An experimental study of the eye-voice span in reading. *Suppl. Educ. Monog.*, 1920, No. 17. Pp. xi+105.
3. ———. Fundamental reading habits: a study of their development. *Suppl. Educ. Monog.*, 1922, No. 21. Pp. xiv + 150.
4. ———. Diagnostic studies in arithmetic. *Suppl. Educ. Monog.*, 1926, No. 30. Pp. 1-46.
5. ———. A laboratory study of the reading of modern foreign languages. New York: Macmillan, 1927. Pp. xii+100.
6. Butsch, R. L. C. Eye movements and the eye-hand span in typewriting. *J. Educ. Psychol.*, 1932, 23, 104-121.
7. Dearborn, W. F. The psychology of reading. *Columbia Univ. Contrib. Phil. & Psychol.*, 1906, 14, No. 4. Pp. 134.
8. Dockeray, F. C. The span of vision in reading and the legibility of letters. *J. Educ. Psychol.*, 1, 123-131.
9. Dodge, R. Visual perception during eye movements. *Psychol. Rev.*, 1900, 7, 454-465.
10. ———. The illusion of clear vision during eye movements. *Psychol. Bull.*, 1905, 2, 193-199.
11. ———. An experimental study of visual fixation. *Psychol. Rev., Monog. Suppl.* 1907, 8, No. 35. Pp. iv+95.
12. Erdmann, B., & Dodge, R. Psychologische Untersuchungen über das Lesen auf experimenteller Grundlage. Halle: Neimeyer, 1898. Pp. viii+360.
13. Eurich, A. Four types of examination compared and evaluated. *J. Educ. Psychol.*, 1931, 22, 268-278.
14. Gray, C. T. Types of reading ability as exhibited through tests and laboratory experiments. *Suppl. Educ. Monog.*, 1917, No. 5. Pp. xiii+196.
15. Hamilton, F. M. The perceptual factors in reading. *Columbia Univ. Contrib. Phil. & Psychol.*, 1907, 17, No. 1. *Arch. Psychol.*, 1906, 1, No. 9. Pp. 57.
16. Huey, E. B. The psychology and pedagogy of reading. New York: 1909. Pp. xvi+470.

17. Jacobsen, O. An analysis of eye-movements in reading music. Unpublished thesis, Univ. Chicago, 1926.

18. JUDD, C. H., & BUSWELL, G. F. Silent reading: a study of the various types. *Suppl. Educ. Monog.*, 1922, No. 23. Pp. xiii+160.

19. LANDOLT, —. Nouvelles recherches sur la physiologie des mouvements des yeux. *Arch. d'ophtalmol.*, 1891, 11, 385-395.

20. MILES, W. R., & SHEN, E. Photographic recording of eye movements in the reading of Chinese in vertical and horizontal axes; method and preliminary results. *J. Exper. Psychol.*, 1925, 8, 344-362.

21. REBERT, G. N. A laboratory study of the reading of familiar numerals. *J. Educ. Psychol.*, 1932, 23, 35-45.

22. ————. A laboratory study of the reading of familiar formulas. *J. Educ. Psychol.*, 1932, 23, 192-203.

23. RUCH, G. M., *et al.* Objective examination methods in the social studies. Chicago; Scott, Foresman, 1926. Pp. 116. Also see *J. Educ. Psychol.*, 1926, 17, 368-375.

24. RUCH, G. M. The objective or new-type examination. New York: Scott, Foresman, 1929. Pp. x+478.

25. RUEDIGER, W. C. The field of distinct vision. *Columbia Univ. Contrib. Phil. & Psychol., & Educ.*, 1907, 15, No. 5. Pp. 69.

26. SCHMIDT, W. A. An experimental study of the psychology of reading. *Suppl. Educ. Monog.*, 1917, No. 2. Pp. iv+192.

27. SHEN, E. An analysis of eye movements in the reading of Chinese. *J. Exper. Psychol.*, 1927, 10, 158-182.

28. STUMPF, N. F. Listening versus reading method in the true-false examination. *J. Appl. Psychol.*, 1931, 15, 555-562.

29. TERRY, P. W. How numerals are read: an experimental study of the reading of isolated numerals and numerals in arithmetic problems. *Suppl. Educ. Monog.*, 1922, No. 18. Pp. xiv+110.

30. THORNDIKE, E. L. The measurement of intelligence. New York; Columbia Univ., Teach. Coll., Bur. Educ., 1927. Pp. xxvi+616.

31. TINKER, M. A. A photographic study of eye movements in reading formulae. *Genet. Psychol. Monog.*, 1928, 3, 68-182.

32. ————. Eye movements duration, pause duration, and reading time. *Psychol. Rev.*, 1928, 35, 385-397.

33. ———. Numerals versus words for efficiency in reading. *J. Appl. Psychol.*, 1928, 12, 190-199.

34. ———. Photographic measures of reading ability. *J. Educ. Psychol.*, 1929, 20, 184-191.

35. ———. Apparatus for recording eye movements. *Amer. J. Psychol.*, 1931, 43, 115-118.

36. ———. The relation of speed to comprehension in reading. *School & Soc.*, 1932, 36, 158-160.

37. VERNON, M. D. The experimental study of reading. Cambridge, England: Univ. Press, 1931. Pp. 190.

UNE ÉTUDE DES MOUVEMENTS DES YEUX DANS LA LECTURE DE QUESTIONS OBJECTIVES D'EXAMEN

(Résumé)

Des enregistrements photographiques des mouvements des yeux, lesquels ont donné des mesures de la fréquence des fixations, de la durée des pauses, et de l'ordre de succession des changements saccadés, ont été obtenus pour 66 étudiants universitaires de deuxième année pendant qu'ils ont lu un *paragraphe de prose scientifique et six types divers de questions objectives* d'examen. Les contrôles ont été faits de sorte que la variation (1) du contenu de la matière, (2) des différences de difficulté entre l'un type et le suivant, (3) des dissimilitudes dans l'apparence typographique, (4) des avantages possibles de l'ordre de présentation, (5) du choix des sujets, et (6) du manque de l'uniformité dans les directions ont été éliminés ou du moins on a fait réduire au minimum leur influence comme variables possibles. Dans ces conditions il paraît justifiable d'attribuer des différences quelconques observées dans la forme des mouvements des yeux pour les divers types des questions, et en comparaison avec *la prose, à l'effet de* la forme des questions.

Les *résultats* d'une *étude quantitative des enregistrements sont les suivants*:

1. Les types des questions rangés en ordre de l'augmentation de la durée de la perception sont les suivants: vrai-faux, complétion, choix multiple, analogie, et mot faux. Cet ordre ne varie que peu quand on les range selon le nombre des fixations, la durée des pauses, ou les régressions. Le variable le moins important des mesures des mouvements des yeux est la durée des pauses. Dans toutes les mesures les moyennes pour les questions dépassent celles pour la prose, de montants divers.

2. A peu près toutes les corrélations entre la prose et les diverses questions, à l'exception de la durée des pauses, sont presque nulles; en contraste avec ceci, les intercorrélations des questions sont relativement assez élevées. Cette différence reflète probablement une modification de l'attitude ou disposition laquelle reste assez semblable pour les diverses questions mais diffère de celle qui est caractéristique de la lecture de la prose.

3. Il n'existe nulle relation constante entre la difficulté des parties (cinq niveaux ont été représentés dans chaque type) et les mesures des mouvements des yeux.

L'analyse qualitative des enregistrements montre que les deux différences les plus générales qui distinguent la lecture des questions objectives et la lecture de la prose sont la suite plus analytique des mouvements des yeux et la variété frappante des procédés employés dans la lecture des questions. Dans chaque exemple, cependant, le moyen adopté pour un type spécifique semble suffisant à la situation. Malgré l'exercice constant avec la prose, cette variété de procédés apparement efficaces et les caractéristiques spéciales telles que les longues pauses sur les complétions et la lecture partielle des questions de choix multiple et d'analogie indiquent que les processus oculomoteurs sont subordonnés et flexibles dans leurs réponses aux processus centraux de l'assimilation et de l'association.

FRANDSEN

EIN AUGENBEWEGUNGSSTUDIUM VON OBJEKTIVEN EXAMENFRAGEN

(Referat)

Photographische Aufnahmen der Augenbewegungen, die Werte für Häufigkeit der Fixierungen, Dauer der Zwischenzeiten, und Reihenfolge von Ruckverscheibungen wurden von 56 Universitatsstudenten beim Durchlesen eines Abschnittes einer wissenschaftlichen Abhandlung und sechs verschiedener Arten von objektiven Examenfragen erhalten. Die Kontrollen waren solcher Art, dass die Variation (1) des Inhalts des Materials, (2) Unterschiede der Schwierigkeiten zwischen den verschiedenen Arten, (3) Unähnlichkeiten in der Druckaufmachung, (4) mögliche Vorteile der durch die Ordnung entstandene Darstellung, (5) Auswahl der Vpn., und (6) der Mangel an Einförmigkeit der Richtungen wurden ausgeschalten, oder wenigstens als mögliche Variablen auf ein Minimum reduziert. Unter diesen Umständen scheint es berechtigt zu sein, der Wirkung der Form der Fragen irgendwelche Unterschiede, die in der Gestalt der Augenbewegungen bei der verschiedenen Arten der Fragen beobachtet wurden, und im Vergleich zur Prosa, zuzuschreiben.

Die Ergebnisse des quantitativen Studiums sind folgende:

1. Die Typen der Fragen, die in eine Reihe von zunehmender Wahrnehmungszeit eingeordnet wurden, sind folgende: wahr-falsch, Ergänzung, vielfache Wahl, Analogie, und falsches Wort. Diese Einordnung variiert nur wenig, wenn sie nach der Anzahl der Fixierungen, Dauer der Zwischenzeiten, oder Rückkehr eingeordnet wird. Die geringste Variable der Augenbewegung, die gemessen wurde, ist die Dauer der Pausen. Bei allen Messungen übertreffen die Durchschnittswerte für die Fragen diejenigen für die Prosa schwankend.

2. Fast alle Korrelationen zwischen Prosa und den verschiedenen Fragen, die Pausendauer ausgenommen, nähren sich Null; im Gegensatz dazu sind die Fragenkorrelationen verhältnissmässig hoch. Dieser Unterschied spiegelt wahrscheinlich eine Veränderung der Attitüde oder Einstellung wieder, welche für die verschiedenen Fragen ziemlich gleich bleibt, aber von jener Eigentümlichkeit des Prosalesens verschieden ist.

3. Zwischen der Schwierigkeit der Gegenstände (fünf Schichten wurden bei jeder Art gegeben) und den Augenbewegungswerten war kein gleichmässiges Verhältnis.

Eine qualitative Analyse der Ergebnisse zeigt, dass die zwei allgemeinsten Unterschiede, die das Lesen von objektiven Fragen von dem Prosalesen unterscheiden, sind die analytischere Reihenfolge der Augenbewegungen und die auffallende Verschiedenheit der Verfahren, die beim Lesen der Fragen gebraucht wurden. In jedem Fall aber scheint die Methode, die für einen gewissen Typus angenommen wird, der Situation geeignet zu sein. Dennoch bei stetigen Uebung mit Prosa beweist diese Art scheinbaren wirkungsvollen Verfahrens und solche Sonderzüge, wie lange Pausen bei Ergänzungen und das Teillesen der vielfachen Wahl und Analogiefragen, dass die oculomotorischen Vorgänge untergeordnet und biegsam in ihrer Wirkung im Vergleich zu den zentralen Vorgängen der Assimilation und Assoziation sind.

FRANDSEN

$7.00 per volume
Single numbers $2.00

MONTHLY
Two volumes per year

September, 1934
Volume XVI, No. 3

GENETIC PSYCHOLOGY MONOGRAPHS

Child Behavior, Animal Behavior, and Comparative Psychology

Worcester, Massachusetts

Entered as second-class matter December 1, 1925, at the post-office at Worcester, Mass., under Act of March 3, 1879

$7.00 per volume
Single numbers $2.00

MONTHLY
Two volumes per year

September, 1394
Volume XVI, No. 3

GENETIC PSYCHOLOGY MONOGRAPHS

Child Behavior, Animal Behavior, and Comparative Psychology

AN EXPERIMENTAL STUDY OF CONSTITUTIONAL TYPES*

From the Department of Psychology of Columbia University

By

OTTO KLINEBERG, S. E. ASCH, AND HELEN BLOCK

*Accepted for publication by C. J. Warden of the Editorial Board and received in the Editorial Office, May 7, 1934.

Worcester, Massachusetts

Entered as second-class matter December 1, 1925, at the post-office at Worcester, Mass., under Act of March 3, 1879.

ACKNOWLEDGMENTS

Acknowledgment is gratefully made to Professor J. P. Turner of Brooklyn College for making possible the experiments reported in Chapter II; to Dean Virginia C. Gildersleeve of Barnard College for permission to experiment on Barnard students; to Professor Agnes R. Wayman of Barnard College for the use of anthropometric records; to Dr. Austin B. Wood and Mr. Everett G. Brundage of Brooklyn College for their cooperation in giving tests to their classes; to Florence Miller Asch for her assistance in the construction of the tests and in the preparation of the manuscript; to Mr. Joseph Goldberg for his assistance in the scoring of the tests and the treatment of the results; and to the students of Brooklyn College and Barnard College for their kindness in serving as subjects in this study.

OTTO KLINEBERG
S. E. ASCH
HELEN BLOCK

COLUMBIA UNIVERSITY
NEW YORK CITY

CONTENTS

I

THE PROBLEM

A. Introduction

Constitution may be defined as the totality of individual physical characteristics. These characteristics, of course, vary tremendously, and there have been many attempts to introduce some order into this variability by means of a classification into "types." Probably the first typology was that of Hippocrates, who spoke of a "habitus apoplecticus" and "habitus phthisicus," bodily constitutions which predisposed to circulatory and respiratory diseases, respectively. Since his time there have been a great many other classifications, for the most part inspired by a similar clinical interest, and in many cases leading to somewhat similar conclusions. For historical reviews, see Weidenreich (42), Wertheimer and Hesketh (44), and Saller (33).

In recent years the problem of constitutional types has become the concern of the psychologist, largely as the result of the work of Kretschmer (17, 18), who claims to have discovered an intimate relationship between a certain type of physique and a certain type of mental disease. More specifically, he has suggested on the basis of a statistical study of clinical material that manic-depressive psychosis is likely to be found in association with the "pyknic" constitution, and schizophrenia or dementia praecox, with the "leptosome," and somewhat less frequently, with the "athletic" and "dysplastic" make-up.

The constitutional type of an individual is recognized by a process of diagnosis in which qualitative as well as quantitative observations play a part. Anthropometric measurement is of course important, but it must be interpreted in relation to the whole constitution. When a subject is examined, says Kretschmer, "kein Haar auf seinem Haupt und keine Variante seiner Nasenspitze ist uns gleichgültig" (17, p. 44). A leptosome, for example, is to be diagnosed not only by means of indices which take into account his shoulder breadth or his chest circumference, but also by the fact that in typical cases he has an angular profile, a long nose and a short chin, narrow, tapering hands and feet, etc. Not all leptosomes have all of these characteristics; the type is the perfect case, not the average, and the classification is determined not by the most frequent, but by the most "beautiful" examples. "Daraus ergiebt sich, dass sich unsere Typenbeschreibung, nicht nach den häufigsten, sondern nach den schönsten Fällen richtet" (17, p. 16). The classical cases are exceptions.

This approach of Kretschmer's makes it somewhat difficult to deal with his material in strict statistical fashion. In general leptosomes and pyknics differ, but it may be possible for an individual to have some pyknic characteristics and yet be diagnosed a leptosome. As a result, a certain amount of subjectivity is introduced into the diagnosis, and it is not impossible that two observers may classify the same individual somewhat differently. This situation is perhaps comparable to the problem of diagnosing, let us say, a

case of appendicitis which does not show all the typical or classical features; the diagnosis can usually be made, but with difficulty, and it not infrequently happens that during the operation the infection is found not to reside in the appendix after all. Kretschmer's clinical approach and his insistence on the "Gesamtkonstitution," the total constitutional picture, has made it extremely difficult to check his results by purely objective standards.

Granting all this, the fact remains that Kretschmer's type theory has stimulated a great deal of study and research in Germany and other European countries, although it has had relatively little influence in America. His original application of the theory to psychotic conditions has been extended to include "normal" or healthy individuals, and the methods of experimental psychology have been employed by Kretschmer's followers to justify this extension. The experiments are ingenious and the results striking, but there are important defects in the procedure, particularly in the choice of subjects, and in the lack of adequate statistical controls. The present study is an attempt to discover whether these type differences still appear under more rigid experimental conditions.

The *leptosome* habitus (formerly referred to as *asthenic*) is characterized by a deficiency in volume combined with relatively tall stature. The thorax is long and flat, the shoulders narrow and the trunk cylindrical. The legs and arms are long and thin. The facial profile is irregular, with a long nose and small chin. The *pyknic* has a thickset, barrel-shaped trunk,

a broad chest and large body cavities, with a definite disposition to plumpness. The neck is short and full with the head set forward on smoothly rounded shoulders. The hands are small and broad, with short fingers, the face is shaped like a pentagon, or shield, with a tendency toward a flat or concave profile. The *athletic* type has strongly developed bones and muscles, large shoulders and a tapering trunk, long limbs and large hands and feet. The face is described as an elongated egg form. The *dysplastic* type is really a defective type, showing a lack of development in one or more organs of the body.

Of these types the leptosome and the pyknic appear to be most important, and will be referred to exclusively in what follows. They correspond very closely to the constitutional types recognized and described by other investigators; for example, to Viola's *microsplanchnic* and *macrosplanchnic*; Pende's *longilineus* and *brevilineus*; Stockard's *linear* and *lateral*; Weidenreich's *leptosome* and *eurysome*. There seems an almost universal tendency on the part of typologists to distinguish between what Weidenreich has called the *long-thin* and the *short-thick* body types, and to look for their clinical or psychological correlates.

Kretschmer's starting-point was his observation that there was a relation between constitutional type and two of the most important diseases in the Kraepelinian classification, namely schizophrenia or dementia praecox, and manic-depressive or circular insanity. In 1921 (18) he reported that among 260 patients studied, 70.3% of the schizophrenics were of leptosome build,

and only 2.9% pyknic, while of the manic-depressive patients 84.7% were pyknic and only 10.6% leptosome. This finding stimulated a considerable amount of investigation in other laboratories, to such an extent that in the 1931 edition of *Körperbau und Charakter* (17) he was able to summarize the results of several investigators, including over 4000 cases, with the following results:

	Schizophrenic	Manic-Depressive
Pyknic and mixed pyknic	12.8%	66.7%
Leptosome and athletic	66.0	23.6
Dysplastic	11.3	0.4
Atypical	9.9	9.3

Most of the studies in this field have confirmed Kretschmer's theory, but criticisms have by no means been wanting. [See especially Möllenhoff (24), Kolle (15, 16), Weidenreich (42), and Garvey (9)]. These have centered mainly on the age factor (both pyknicity and manic-depressive insanity tend to be found rather late in life), social and economic factors (the leptosome habitus and the schizothyme temperament may be relatively frequent in the more favorable and more "exclusive" environments), lack of normal controls (the proportion of physical types in schizophrenia appears to be very similar to that in the normal population), the subjectivity and the consequent probable bias inherent in the constitutional diagnosis, and the unsatisfactory nature of the Kraepelinian classification of the psychoses.

Kretschmer believes that the correspondence between physique and character holds for normal and borderline, as well as for psychotic cases. The borderline

subjects, who may be regarded as pre-psychotic personalities, fluctuate between health and disease, and can be studied in the case histories of the early life of psychotics. They are classified as *cycloid* and *schizoid*; the cycloids are usually pyknic, and the schizoids usually leptosome. Cycloids are sociable, fond of people, adaptable, realistic, practical. Even when they are unhappy, they like people about them to whom they can tell their troubles. They work hard, but are not especially ambitious. Schizoids on the other hand are not very sociable, or sociable only in small restricted circles. They are very difficult to know, because they live in themselves so much; they are "autistic." They are exclusive and "aristocratic," even in peasant circles; there is always a glass cage between them and the universe.

The normal biotypes are the *cyclothyme* and the *schizothyme*, again related to the pyknic and leptosome physique respectively. There are several subvarieties: the cyclothyme may be talkative and gay, or quietly humorous, or a comfortable Epicurean, or quietly friendly and sociable, or successful and practical. The schizothyme may be aristocratic and sensitive or dry and empty, an unworldly idealist or a cool masterful egoist. These two types differ not only in their social reactions, their choice of an occupation, their susceptibility to one or another form of maladjustment, but also in their behavior in the psychological laboratory.

B. THE EXPERIMENTAL APPROACH

The experimental studies in this field, many of which were made by Kretschmer's students and associates in Germany, cover a wide range of psychological activities, and are regarded by Kretschmer as furnishing a valuable confirmation of his theory (see 17, 39, 43). A brief summary follows.

In *perception span,* for example, pyknics are superior to leptosomes, and circulars (manic-depressives) superior to schizophrenics. [*Van der Horst* (40), *Vollmer* (19), Enke (3), Kibler (14)]. In a *reaction time* experiment, pyknics and cyclothymes are much more disturbed by distractions than are leptosomes and schizothymes; for the former group there is a greater difference between their reaction times with and without distraction. [Kibler (14), Van der Horst (40)]. These results would corroborate Kretschmer's view that pyknics and circulars are more objective, and more attentive to external stimuli.

Leptosomes excel in "Spaltungsfähigkeit" or "cleavage capacity," the capacity to split the stream of consciousness so that parts of it function separately (although simultaneously) and for disparate purposes [See Kretschmer (17)]. They can "abstract" better.

When Kibler (14) exposed tachistoscopically cards containing colored nonsense syllables in varying positions, and asked his subjects to note the color and position, the leptosomes were much more successful, although the pyknics remembered more about those elements which they were not asked to remember. They

showed more "incidental memory," but were less successful in "abstraction," in the orderly separation of elements. Enke (3) measured "Spaltungsfähigkeit" by presenting to his subjects a series of colored squares, with directions to count in passing the number of squares of each color in the series. The task required the ability to keep the separate impressions distinct from each other, and the leptosomes were again superior. This result was corroborated by Enke and Heising (6). Enke (3) studied this same cleavage phenomenon by another method. He exposed tachistoscopically very long, hybrid words (for example, "Badevereinsmarke"), each word exposed ten times for .3 of a second each time, the subjects being instructed to report what they saw after each exposure. Leptosomes tended to reconstruct the word in an analytical, systematic manner, building it up letter by letter with each successive exposure. Pyknics responded synthetically, obtaining a general impression of the word, and correcting this impression after subsequent exposures. They made many more mistakes than did the leptosomes, who were more orderly, and who abstracted better from the total exposure the elements for which they were to look. [See also Lutz (19), Vollmer (19)].

Pyknics are more sensitive to colors, leptosomes to forms. Scholl (35) found that when he showed a colored figure to his subjects, and asked them to pick this figure out of a larger group of various colors and forms, schizoids tended to identify the figure according to its form, cycloids according to its color. Enke (3),

Dambach (19), and Lutz (19) found these differences to hold for pyknics and leptosomes. Pyknics also have a greater capacity for color blending; when red and green discs are combined on a color wheel, flicker disappears sooner, and grey appears more quickly, in the case of the pyknic observers, leptosomes requiring a faster rate of revolution. [Van der Horst (40), Wiersma (46), Kibler (14)].

The greater color reactivity of the pyknics also plays a part in their responses in the Rorschach test (32). This test, consisting of a series of irregularly-formed ink blots, some of them in color, has been used by psychiatrists as a personality test, on the theory that the responses reflect the interest and the mental organization of the subject. Rorschach gives instructions for diagnosing extravert ("extratensiv") and introvert ("intratensiv") personalities on the basis of the responses, color responses in general suggesting extraversion, movement responses introversion. Munz (25) found that pyknics gave more color responses, responded more to brightness qualities, to objects and landscapes, and gave more "total" responses, i.e., to the whole figure at once. Leptosomes gave more movement responses—faces, dancers, dream-figures, unreal objects—and many more "detail" responses, i.e., to one part of the ink-blot at a time. There were more "extraverts" among the pyknics, and more "introverts" among the leptosomes. Enke (2) obtained substantially the same results.

Personality differences have also been studied by means of a self-diagnosis by pyknics and leptosomes

[Van der Horst (40), Kibler (14)]. The subjects were given two lists of characteristics to choose from, one list based on Kretschmer's description of the cyclothyme, the other on that of the schizothyme, and were asked to choose the list which described them more accurately. Almost all the pyknics chose the cyclothyme list, and the large majority of leptosomes chose the schizothyme list.

In this country Naccarati and Garrett (27) found a slight positive relationship between macrosplanchny (roughly identical with "pyknicity") and emotional instability as measured by the Woodworth Personal Data sheet. A follow-up study by Garrett and Kellogg (8) failed, however, to corroborate this finding.

The word-association test shows a greater tendency for leptosomes to give meaningless associations, or associations difficult for the experimenter to understand (autism?), while pyknics give rather more predicative or "feeling" associations. Leptosomes tend to perseverate, frequently giving the same response to dissimilar words [Van der Horst (40), Bayer (19)]. They appear to have their thought processes concentrated around a central point, while the pyknics have true "chain" associations. Murphy (26), however, failed to find any clear differentiation in the types of word association given by manic-depressives and schizophrenics.

Studying "affectivity" by means of the psychogalvanic reflex, Enke (5) found that when the subjects were first placed in the experimental situation it took a longer time for leptosomes to reach a "resting" level.

Leptosomes also reacted more violently to a variety of sensory stimuli (i.e., odors and loud sounds), and when engaged in mental work (arithmetic), but pyknics showed more marked responses to pain stimuli. Enke says this indicates the greater sensitivity of pyknics to "vital feeling" as against the insensitivity of schizoids (cp. catatonic symptoms in schizophrenia).

On the motor side, leptosomes tap faster when told to tap at their natural or most comfortable rate [Van der Horst (40), Enke (4)]. In motor activities in which many muscles come into play, pyknics are superior; their movements are more rounded and flowing [Enke (4), Oseretsky (28), Ossipowa and Ssucharewa (29)]. Very fine and delicate movements are, however, carried out better by leptosomes. Fatigue comes on gradually in the case of pyknics, suddenly in that of leptosomes. The handwriting of the two groups also differs [Jislin (12), Haarer (10)]. Pyknics write a flowing hand, the letters connected, the writing straight and of medium height, the letters very similar to each other; there is a similarity in the writing of most pyknics. Leptosomes differ much more from each other, but their writing is usually angular, with uniform pressure, the letters often separated, the writing frequently infantile and uncertain, with many stereotyped curves and figures.

As far as intelligence is concerned, Kretschmer does not lay any stress on possible differences between constitutional types. Naccarati (27), however, suggested that intelligence was slightly but positively correlated with microsplanchny (corresponding to the leptosome

habitus, but Garrett and Kellogg (8) repeated the study with negative results. Mohr and Gundlach (23), in a study made at Joliet prison, found leptosomes to be superior to pyknics in Army Alpha, but this may be accounted for by a difference in the background of the two groups. It may be added that their other tests tended on the whole to corroborate the findings of the German investigators, but the differences were small, and the relationship between physique and "temperament" at best a slight one.

The more important general conclusions arrived at by the German investigators may be summarized as follows. Pyknics (cyclothymes) as contrasted with leptosomes (schizothymes) have a greater perception span, are more distractible, have less power to abstract, less "cleavage capacity" or "Spaltungsfähigkeit," better incidental memory, respond "synthetically" (leptosomes "analytically") in the case of a difficult perception, are more sensitive to colors (leptosomes to forms), see colors as blends more quickly, respond as "extraverts" to the Rorschach test (leptosomes, "introverts"), diagnose themselves as cyclothyme (leptosomes, schizothyme), give more chain associations in the word association test (leptosomes more perseveration and more meaningless associations), show less marked psychogalvanic reactions except to pain stimuli, and are superior in motor tasks except those which require fine and delicate movements. The handwriting differences are not clear, and the differences in intelligence negligible.

C. CRITICISMS OF TYPES

Kretschmer's theory of types is important precisely because it is a theory. It offers highly interesting and challenging suggestions for the study of the organization of mental and emotional activities along causal lines, in contrast with the prevailing more random and empirical procedures. It also offers the hope of a *systematic* account of individual differences and of the relation between mental traits. In spite of this, the theory has been relatively neglected by American psychologists.

The reasons for this neglect are not hard to find. Kretschmer's theory of types—and for that matter every type theory—is the outgrowth of qualitative analysis; although it supplements its assertions by experiments, these are considered secondary and supplementary to the main thesis.

Mental measurement in America, on the other hand, has traditionally stressed the quantitative and statistical approach to the problems of individual and group differences. In addition, however, to the differences in intellectual and historical backgrounds prompting these two "types" of approach, there is a more specific reason for the distrust of type psychologies in America. The reason is an empirical one. Quantitative measurement, and specifically the normal distribution of abilities, is held to have furnished definite proof of the presence of *one* type, namely the average, whereas type psychologies seem to single out a few extreme and rare cases, emphasize their contrasts, and arbitrarily constitute them into types.

A clearer understanding of the objection to type theories may be obtained through a detailed analysis of a representative and comprehensive criticism. Probably the best and most influential of these criticisms is that of Thorndike, vigorously elaborated in his *Educational Psychology* (38). Thorndike asserts that type formulations are in conflict with experimental evidence on the following grounds:

(1) Differences in ability grade into each other continuously.

(2) Only one central tendency is generally found in distributions of performance scores. On the other hand, no single tendency would be present with more than one type.

(3) Multi-type theories imply the existence of multimodal distributions. No such distributions have been satisfactorily demonstrated.

(4) "Pure" types are rare. They are extreme and atypical members of their groups, and more plausibly to be conceived as deviates from the norm. The criticism is here implied that conclusions based on "pure" types could not be generalized to the bulk of the population.

(5) Negative correlations between mental traits are relatively rare. They are demanded in type theories.

As one considers the various objections marshalled against multitype theories the interesting situation is presented in which the critics ignore the assumptions on which type theories are based. This seems to be the clue to the understanding of type criticisms. The types that are attacked are statistical types, but type concepts are not statistical: they are neither contradicted by the presence of normal distributions of ability nor do they require multimodal distributions.

The confusion seems to be produced by the unac-

knowledged conversion of the central tendency from a statistical reference point into an experimental reference point. It is undoubtedly convenient to treat all points on a curve of ability as deviations from the mean. It may be equally convenient and "true," when the purpose is scientific explanation, to regard the central tendency as a deviation from both extremes. This is precisely what the geneticist does when he explains his "mixed" and numerically preponderant mongrel groups in terms of extremes of "pure" and rare types.

That the notion of a single type is an *inference* from the fact of normal distributions of ability is recognized at one point by Thorndike, as is evident in the following statement: "By the theory of a single type, one make-up can be conceived such that from it all individuals would differ less than they would from any other make-up, and such that the greater the divergence, the rarer they would be." This statement is a powerful argument for a single-type theory of mental abilities. Thorndike believes, however, that this is the only possible inference, for he goes on to say that "by the theory of multiple types, no such single true central tendency would exist."

We cannot agree with the latter statement, for as has been indicated above, "pure" types may in successive generations act so as to produce a numerically preponderant central tendency. It seems clear, therefore, that the fact of the normal distribution of ability (or the absence of multimodal distributions), the presence of single central tendencies, and the rarity of extreme cases are not conclusive objections to multiple-type theories.

Up to this point our considerations have centered on the distributions of *single* abilities. Type theories are, however, fundamentally theories of the *organization* of mental abilities, and the term type refers to a necessary, or causally connected, organization between different traits. It is by examining the relationships between abilities that we may be able to state the conditions which a theory of types must meet if it is to be scientifically established.

What lends distinctiveness to a theory of types is not the assertion that type *A* differs from type *B* in function *X*. The differences between the types are *general* and extend over a wide range of mental activities. Thus, type *A* may be superior in functions *X* and *Y*, inferior in function *Z*, etc. On the basis of these statements, we may indicate one preliminary condition for the proof of a type theory: the functions which distinguish between two types must conform to the statistical requirements of a *mental trait*, or *group factor*. The proof of a group factor has been stated by Kelley (13). It consists of the demonstration of a central bond that relates the abilities in question, and is independent of other abilities or traits.

Group factors do not, however, establish the presence of types. It is the assumption of all correlational studies that the results they find are true of their groups as a whole. It is the assumption of group-factor studies that their results would remain unaltered (apart from the influence of the number of cases and of chance errors of sampling) if they were based upon any portion of their group selected by means of an external

criterion (such as physique). It is here that the differences between group factors and types become apparent. For according to the theory of types, correlational results will change radically with the change in the composition of the groups. Suppose that type A is superior to type B in functions X and Y. Then a correlation between X and Y obtained with a population composed of both types is bound to be high and positive. Conversely, suppose that type A is superior to type B in function X and inferior to type B in function Z. The correlation between X and Z, when based on a population consisting of both types, will be high and negative. Both correlations will become, however, drastically reduced when the population consists of only one type. We are now ready to state the necessary condition for the proof of the presence of types: the correlations between abilities within a *single* type must be significantly smaller than the correlations between abilities for both types taken together.

In this connection, the criticism advanced by Thorndike and others concerning the relative absence of negative correlations becomes important. A type theory seems to demand negative correlations between functions, when one type is superior in one function and inferior in the other. This criticism is significant. The qualification should be made, however, that the possibility of negative correlations has not been disproven and that type theories deserve attention, if for no other reason, because they might suggest the locus of such negative relationships.

It has been the purpose of this discussion to show that

the concept of types can be placed on a sound scientific footing and that strict conditions for its proof may be developed. We have tried to show that both the concept of a single and of multiple types are *inferences* from the facts of measurement, and that both are plausible scientific hypotheses. It has been our hope to conciliate two divergent lines of thought, first by establishing the claim of the concept of types to a position of importance, and then by pointing out that its final verification can be accomplished only by the use of quantitative and statistical methods.

II

THE FIRST STUDY

A. Procedure

1. *The Subjects.* The necessity for a considerable degree of homogeneity with regard to age, sex, race, educational and social background for subjects in a type study will become apparent in subsequent discussion. The subjects of this study met the conditions of homogenity in a highly satisfactory manner. They were selected from a group of 153 students in a course in general psychology, in Brooklyn College. All the subjects were males. All, with the exception of 4, were Jewish. Their average age was 19 years and 9 months, with a standard deviation of 17 months. The great majority had received their education in the school system of New York City. In addition, they were relatively homogeneous with regard to occupational background. Approximately 75% of the subjects came from homes of shopkeepers, merchants, salesmen, etc., while the bulk of the remainder had parents who were skilled laborers. The reader will find more detailed descriptions of the background of groups from the same college system in the studies of Schneck (34), Anastasi (1), and Smith (36).

2. *Physical Measurements.* The first task was to select accurately subjects for the leptosome and pyknic categories. No completely satisfactory method exists for such classification. The important difficulty is the fact upon which all investigators in this field agree—that there are large numbers of individuals who possess

traits credited to both types. Thus Kretschmer is very frequently led to speak of individuals who are "predominantly" leptosome or pyknic. To use his own words, "we could reel off here, and with other types, innumerable mixtures of such a kind: there is absolutely no single criterion which cannot be varied by and combined with marks of another type." This situation was fully confirmed with our subjects. It is apparent therefore that separation of subjects into physical groups must be empirical and to some extent arbitrary, pending the development of more reliable norms than we now possess.

The physical traits measured were the following: standing height, sitting height, weight, shoulder breadth, and chest circumference. All measurements were taken in accordance with the standard procedures described by Martin (22). The chest circumference was measured with a cm. steel tape; the shoulder breadth was measured with a large pair of calipers.

On the basis of these measures, a number of indices were computed which have been used in the past by a large number of investigators, to differentiate between the physical categories. These were as follows:

$$(1)\ \frac{\text{Standing Height (in cms)}}{\text{Weight (in kgs)}}$$

$$(2)\ \frac{\text{Standing Ht. (in cms)}}{\text{Sitting Ht. (in cms)}}$$

$$(3)\ \frac{(\text{Standing Ht.})^2}{\text{Sitting Ht.} \times \text{Weight}}$$

$$(4)\ \text{Pignet Index} = (\text{Weight} + \text{Chest Circumference}) - \text{Height}$$

$$(5)\ \text{Chest — Shoulder Index} = \frac{\text{Shoulder Breadth}}{\text{Chest Circumference}}$$

In view of the absence of clear-cut criteria, three different procedures were employed to separate the subjects into the main physical groups.

The first method of classification consisted of the experimenter's estimate or "diagnosis" of the subject's physique, based on general appearance. All estimates were made with the subjects completely stripped. These estimates were arrived at before objective measurements were taken and may consequently be regarded as independent of the objective measurements. The experimenter's estimate took into account not only the impression produced by the relationships of the physical features, but also included a set of features not readily accessible to quantitative measurement, but regarded by Kretschmer as having diagnostic significance, such as the contour of the face and neck, the firmness of the musculature, the amount of adiposity, the shape of the hands and feet, the color and turgor of the skin, etc. This method we will designate as the method of *Diagnosis*.

The distributions of the physical measures of the leptosomes and pyknics, as classified by the method of diagnosis, showed large non-overlapping areas which diverged from each other in the expected directions. These distributions served as the starting point for the second mode of classification, i.e., the method of *Indices*. The boundaries of the intermediate areas where most overlapping occurred were adopted as the limiting values for inclusion in one of the two categories. All values of an index above one boundary were regarded as characteristic of one physical group,

while all values below the other boundary indicated membership in the other group. Subjects falling within the intermediate zone were regarded as "doubtful" and were eliminated from this study. When all of the significant index scores of a subject placed him with one group he was included in that group. This procedure eliminated practically all overlapping between the distributions of the important indices for the leptosomes and pyknics.

While the agreement between the classifications of the two above-mentioned methods was very close, it was not perfect. We therefore adopted a third method of classification, by the formation of a third group consisting of only those subjects who were classified in the same way by *both* the indices and the experimenter's estimate. This classification is presumably more accurate than either of the other two in which the criteria were either the objective measures alone, or the experimenter's estimate alone. We will designate this procedure as the method of *Indices and Diagnosis*. Subsequent comparisons of the performances of leptosomes and pyknics will be made separately for each of these three methods of classification.

3. *The Tests.* In our selection of tests for this investigation, we were guided primarily by the attempt to measure those functions which are regarded by the German investigators as differentiating the leptosome and pyknic types. The following tests were given:

The *Otis Self-Administering Test of Mental Ability* was given under a time limit of 20 minutes. This test was included not for the

purpose of discovering possible differences between the leptosome and pyknic groups—for all the relevant evidence indicates the unlikelihood of such a result (17), but rather to make certain that the two groups were equated in intellectual level. If any significant differences were later found with other tests, it would have been possible to rule out the factor of intelligence as an explanation.

Lecky Individuality Record. This inventory, constructed by Lecky (21), consists of 160 questions, attempting to measure the subject's general attitude toward a number of social and personal problems. In a recent study by Pallister (30) this test was reported to show a general factor, indicative of a negative or withdrawal attitude. The strength of this factor is very highly correlated with the total raw score. The test is divided into eight categories, designated as optimism, habits, nervousness, social adjustment, cooperation, physical well-being, sex attitudes, attitudes toward the family. In addition to the total score, a score of unevenness or inconsistency was obtained. This measure was determined by subtracting from the average score (the total score divided by 8) the scores for each of the eight component parts of the test, and adding the resulting deviations.

General Information. This was a test of general information, devised to measure the *extent* and *uniformity* of the subject's acquaintance with a set of facts drawn from 13 different fields of subject matter. The materials for this test were partly taken from existing achievement tests and partly made up by the authors. The total score was taken as a measure of the extent of information, while the unevenness of the scores was calculated in the same way as the unevenness on the Lecky Individuality Record. The test consisted of 130 questions. No time limit was set for this test.

Digit Memory Span. Two forms of this test were presented; one using the visual method, the second, the auditory method. The number of digits read ranged from 4 to 13, one second being allowed for the reading of each number or for its presentation. Four forms of each test were presented. The score was the average memory span. It will be observed that the auditory method involves the *successive* presentation of *single* numbers, while the visual method involves the *simultaneous* presentation of *all* the numbers. The purpose was to find whether the relative position of two groups would shift at all as the method of presentation was varied.

Spaltungsfähigkeit Test or test of "cleavage capacity." The purpose of this measure, and of those to be subsequently described, was to compare by means of a variety of materials, the ability of the two groups to respond to a large number of stimuli presented simultaneously or successively. According to Kretschmer, the leptosomes are characteristically superior and the pyknics consistently inferior in the performance of such tasks. That this capacity for responding to a multiplicity of factors within a single situation marks *the* fundamental difference between the types is indicated in the following statement from Kretschmer:

> It is an elementary fact of constitution, and is, for example, much greater among individuals of leptosomatic bodily build, than in those of pyknic physique, whatever the experimental situation may be. This cleavage capacity is a root characteristic to which we can trace back a whole series of complicated mental properties which are of fundamental significance in differentiating types of personality and ability, both with normal people and those bearing the stamp of genius. Such mental characteristics are: the inclination to abstract or to concrete thinking, the tendency to analysis or synthesis in reasoning and perception; and, in the affective life, the tendency, or the absence thereof, to build up emotionally-toned complexes. (18, pp. 50-51)

Each trial of the "cleavage capacity" test consisted of a group of letters read to the subjects, each letter being repeated once or more than once during the reading (e.g., a-b-b-a-a-b-). The readings were recorded on a phonograph record observing an interval of two seconds between successive letters. The subjects were instructed to record after the total presentation, on appropriately prepared answer sheets, the number of times each letter was repeated. They were strongly cautioned to avoid counting with any part of the body[1] and in general, to make no use of any "external" aids. The test was varied according to the number of different letters employed in a single reading and according to the length of each reading. The number of different letters read varied from 2 to 4 (*A-B*; *C-L-P*; *D-K-R-T*), while the total number of letters read varied from 6 to 22. A similar procedure was followed in the presentation of lists of

[1]The records of four subjects who were observed making use of aids, such as counting with the fingers, etc. were eliminated.

numbers. The different numbers read were *1-2; 3-5-7;* and *4-6-8-9.* The length of the lists varied, again, from 6 to 22. The scores were stated in terms of errors, and errors were expressed in terms of the difference between the actual number of readings of each letter and the reported number of readings.

Cancellation of Letters. These tests consisted of mimeographed sheets of unrelated letters. The instructions to the subjects were to cancel given letters as rapidly as given compatible with maximum accuracy. The length of each trial was 1 minute; the interval between trials was 30 seconds. After a preliminary set of 5 trials in the cancellation of *a*'s, designed to equalize some of the general differences in past experience and in the adjustment to the task, the following tasks were presented: the cancellation of letters *B-R; D-K-S; F-M-Q-X; C-H-P-T-Z.* The tests became increasingly more difficult as the number of letters to be cancelled increased.

Cancellation of Numbers. A similar procedure was followed with a set of tests involving the cancellation of numbers. There were four trials in the cancellation of each of the following groups of numbers: *1-2; 3-5-7; 2-4-6-8.* The tasks were rotated, and here too they were preceded by five preliminary trials in the cancellation of *O*'s. The purpose, again, was to find, first, whether the groups differed on any of the tests and, second, whether the relative standing of the groups would shift as the number of items to be cancelled during any single trial was increased.

A Test of Incidental Memory. The subjects were presented with a connected but unfamiliar English passage dealing with the social life of the Eskimo. They were under instructions to underline all words containing one or more *a*'s. As soon as a subject completed a task, his time was recorded and he was given a list of 100 questions based on the content of the passage. Each question could be answered in three different ways; as true, false, or uncertain. Scores were expressed in terms of the number of right answers and of wrong answers.

This test, which has generally been used to measure 'logical' memory, or memory for ideas, has not been administered before in this way. It was included in this series in order to test individual memory as well as to determine whether there were any differences between the constitutional groups in their ability to perform the given task without being distracted by the surrounding meaningful material.

All the above eight tests were administered in groups of approximately 20.

B. Results

1. *Physical Results.* The results obtained with the physical measures, including the averages, standard deviations, and the reliability of differences between the pyknic and leptosome groups, are presented in Tables 1, 2, and 3. Reliable differences between the

TABLE 1

Means, Standard Deviations, and Reliability of Differences of Physical Measures of Leptosome and Pyknic Groups when Selected on the Basis of Experimenter's Diagnosis

Measure	Leptosome*	Pyknic**	Difference	D/σd
Height: standing	173.24 ± 5.36	168.72 ± 5.12	4.52	4.10
Height: sitting	90.26 ± 2.82	88.70 ± 3.42	1.56	2.44
Weight	63.20 ± 6.44	69.12 ± 10.16	5.92	3.33
Shoulder breadth	38.23 ± 1.80	37.22 ± 1.91	1.01	2.73
Chest circumference	90.89 ± 4.85	94.85 ± 5.73	3.96	3.67
$\frac{\text{St. Height} \times 100}{\text{Weight}}$	277.10 ± 21.55	252.22 ± 29.35	24.88	4.56
$\frac{\text{St. Height} \times 100}{\text{Sitt. Ht.}}$	192.28 ± 5.16	189.55 ± 4.95	2.73	2.76
Pignet Index	+19.80 ± 9.10	+4.50 ± 11.05	15.30	7.41
$\frac{\text{St. Ht.}}{\text{Wt.}} \times \frac{\text{St. Ht.}}{\text{Sitt. Ht.}} \times 100$	532.1 ± 49.80	467.00 ± 56.40	65.10	6.07
Chest-Shoulder Index	421.1 ± 20.30	403.2 ± 17.50	17.9	4.86

*N=70
**N=40

two groups have been found, for each of the three methods of classification, in the following measures and indices: weight, chest circumference, height-weight ratio, the Pignet index, the chest-shoulder index, and the index $\frac{(\text{standing height})^2}{\text{weight} \times \text{sitting height}}$.

The large differences between the two groups in weight and chest circumference are in accordance with

TABLE 2

MEANS, STANDARD DEVIATIONS, AND RELIABILITY OF DIFFERENCES OF PHYSICAL MEASURES OF LEPTOSOME AND PYKNIC SUBJECTS SELECTED BY OBJECTIVE INDICES ALONE

Measure	Leptosome*	Pyknic**	Difference	$D/\sigma d$
Height: standing	170.96± 5.88	171.52± 6.40	.56	0.50
Height: sitting	89.32± 2.64	90.30± 2.96	.98	1.92
Weight	58.84± 4.00	71.80± 6.56	12.96	13.09
Shoulder breadth	37.78± 1.73	38.55± 1.77	.77	2.41
Chest circumference	88.01± 3.51	95.51± 5.94	7.50	8.43
$\frac{\text{Height} \times 100}{\text{Weight}}$	291.35±15.9	239.75±19.50	51.60	15.78
$\frac{\text{Standing Ht.} \times 100}{\text{Sitting Ht.}}$	192.01± 4.80	190.20± 4.83	1.81	2.06
Pignet index	+24.90± 4.75	+3.85± 8.40	21.05	16.84
$\frac{\text{Ht.}}{\text{Wt.}} \times \frac{\text{St. Ht.}}{\text{Sitt. Ht.}} \times 100$	560.60±35.40	453.50±36.90	107.10	16.11
Chest-Shoulder index	430.3 ±20.30	404.7 ±18.10	25.6	7.23

*N=58
**N=60

TABLE 3

MEANS, STANDARD DEVIATIONS, AND RELIABILITY OF DIFFERENCES OF PHYSICAL MEASURES OF LEPTOSOME AND PYKNIC GROUPS WHEN SELECTED BY INDICES AND DIAGNOSIS

Measure	Leptosome*	Pyknic**	Difference	$D/\sigma d$
Height: standing	172.36± 5.60	168.80± 6.72	3.56	2.41
Height: sitting	89.82± 2.96	89.40± 3.36	0.42	0.57
Weight	59.40± 4.00	72.92± 8.80	13.52	7.91
Shoulder breadth	37.91± 1.74	38.53± 1.76	0.62	1.88
Chest circumference	88.40± 3.36	96.20± 4.62	7.80	7.96
$\frac{\text{St. Height} \times 100}{\text{Weight}}$	290.60±16.80	235.55±21.45	55.05	11.89
$\frac{\text{St. Height} \times 100}{\text{Sitt. Height}}$	192.73± 5.19	189.61± 4.68	3.12	2.74
Pignet index	+22.75± 5.55	—0.15± 9.30	22.90	12.12
$\frac{\text{Ht.}}{\text{Wt.}} \times \frac{\text{St. Ht.}}{\text{Sitt. Ht.}} \times 100$	559.40±35.70	442.10±36.00	117.30	13.93
Chest-Shoulder index	429.1 ±17.80	401.3 ±17.60	27.8	6.70

*N=46
**N=30

the general trend of previous results. [Kretschmer, (17, 18), Van der Horst (40), Mohr and Gundlach (23), etc.]. Of special interest are the large index-value differences between the groups, since they represent relationships between the physical measures. The height-weight ratio, which differentiates significantly between the two groups, is an important though simple index. Mohr and Gundlach (23) report a correlation between it and a much more comprehensive index, namely

$$\frac{\text{chest circ.} + \text{abdomen circ.} + \text{hip circ.} + \text{weight}}{\text{height}}$$

of +.865 for 174 male cases. There was also a correlation of $+.92 \pm .009$ between the same index and the height-weight ratio for 423 women college students.

The Pignet index has differentiated our groups most clearly. This finding is in agreement with that of Farr (7) who found this to be the best of 75 different indices, and who also reports that it correlated best with total objective measurements and general impressions derived from special X-ray and clinical examinations. Of the remaining indices, the chest-shoulder index and the $\frac{(\text{standing height})^2}{\text{sitting height x weight}}$ index have also been valid for the separation of types.

It is of particular interest to note the general agreement of the results obtained with the 3 different procedures. In view of the correspondence of the results of the three methods, it seems reasonable to conclude that they confirm each other's validity.

TABLE 4

AVERAGE MEASUREMENTS OF LEPTOSOME AND PYKNIC GROUPS (WEIGHT IN KGS.; OTHER MEASUREMENTS IN CMS.)

	Kretschmer		Joliet		Van der Horst	Brooklyn College	Kretschmer	Joliet	Van der Horst	Brooklyn College
	Asthenic	Athletic	Asthenic	Athletic	Normal leptosome	Normal leptosome	Pyknic	Pyknic	Normal pyknic	Normal pyknic
Height	168.4	170.0	172.68	174.50	174.5	170.96	167.8	171.74	171.3	171.52
S.D.			7.68	4.34		5.88		4.94		6.40
Weight	50.5	62.9	59.95	67.07	72.6	58.84	68.0	77.64	79.0	71.80
S.D.			5.52			4.00		8.7		6.56
Shoulder breadth	35.5	39.1	37.00	38.62	40.2	37.78	36.9	38.46	37.2	38.55
S.D.			1.83	1.62		1.73		1.78		1.77
Chest circumference	84.1	91.7	85.94	95.54	91.9	88.01	94.5	98.96	97.1	95.51
S.D.			3.53	2.36		3.51		5.02		5.94

Comparisons with results of other investigators. It is important, in studies of this kind, to compare the obtained norms with those given by other investigators for other groups. We are therefore presenting in Table 4 a few comparisons in the measures on which the most important indices were based. It will be observed that there is considerable agreement between the means and standard deviations of the measures of the present and other groups.

Westphal, a co-worker of Kretschmer, has published in a recent study (45) norms for a number of indices based on 439 "pure" physical types selected from 702 men. For these the author gives the curves of distribution and the ranges of the indices. The correspondence between Westphal's values and ours for the most important index, the Pignet, is very close, as is apparent from a comparison of the respective ranges. This is true despite the fact that the figures of Westphal refer rather to the most frequent range, excluding extremes, while our range values represent the difference between minimum and maximum measures. Westphal finds that the range of the pyknics on the Pignet index extends from —10 to +20; our range for pyknics is from —15 to +12. The leptosome range is, according to Westphal, from +10 to +30; our leptosome range extends from +17 to +35. Similarly, close agreement exists between Westphal's norms for the chest-shoulder index and our norms.

2. *Correlations between Physical Measures.* Evidence of a somewhat different character showing the possibility of classifying considerable numbers of in-

dividuals as leptosomes or pyknics is furnished by the size of the intercorrelations between some of the physical measures and indices. These are presented in Table 5. It is noteworthy that our important indices

TABLE 5
INTERCORRELATIONS BETWEEN PHYSICAL MEASURES (N = 140)

	Pignet index	Chest circumference	Height/Weight	Standing height
Chest-Shoulder index	+.514 ±.042		+.430 ±.047	
Height/Weight		—.7734 ± .023		
$\frac{(\text{Sitting height})^2}{\text{Weight} \times \text{Sitt. ht.}}$	+.8840 ±.012			
Sitting height				+.6739 ±.031

show significant correlations with each other (in accordance with the general trend of biometric results), thus pointing to marked relationships among physical traits as a basis for classification of physiques. This finding is in striking contrast to the absence of similar correlations (to be presented shortly) between the mental test scores of this study.

We conclude, as a result of the foregoing discussion, that (*a*) our physical measures are representative of the population from which our group was chosen; (*b*) that the relationships between the measures conform closely to those generally found for groups of college students; (*c*) that the separation of the subjects into two physical categories is valid, in view of the general agreement between our results and those of other type studies; (*d*) that the validity of the classification

is confirmed by the close agreement between three different methods of selection; and (*e*) that the obtained significant correlations between measures and indices lend additional support to the practicability of dividing subjects into different physical categories.

3. *Mental Performances.* In the analysis of the results three separate comparisons between the pyknic and leptosome groups were made in accordance with the three methods of classification used. In each case the averages and standard deviations were computed, and the significance of the differences between the averages was determined.

The results are presented in Tables 6, 7, and 8 and are consistent and unambiguous. As an illustration of our findings, we present in Figures 1 and 2 the distributions for a group of leptosome and pyknic subjects on a physical and mental test. There is practically no overlapping on the physical measure, the Pignet index, while the overlapping on the cancellation test is all but complete. This illustration is characteristic of the trend of our results. They show the absence of any significant differences between the two groups on *any* of the mental tests, for each of the methods of classification used. They fail to confirm, within the limits of the present investigation, the findings of Kretschmer regarding type differences on the mental side.

The same conclusions are furnished by a somewhat different treatment of the results, namely, the correlations between some of the physical indices and some of the mental performance scores. These are presented in Table 9. Without exception, the intercorrelations be-

TABLE 6

MEANS, STANDARD DEVIATIONS, AND RELIABILITY OF DIFFERENCES OF PERFORMANCES OF LEPTOSOME AND PYKNIC GROUPS

(Method of Diagnosis)

Measure	N	*Leptosome* Av.	σ	N	Pyknic Av.	σ	D/σd
Otis	70	45.52	8.06	40	47.93	9.59	1.39
Lecky: total score	70	37.40	19.50	37	36.60	18.40	0.21
Lecky: unevenness score	70	18.36	7.00	37	16.36	5.12	1.68
Information: total score	60	55.92	12.30	33	57.54	10.14	0.68
Information: unevenness score	60	26.43	5.43	33	26.85	5.58	0.35
Auditory span	70	8.33	1.14	40	8.23	1.14	0.42
Visual span	60	9.25	1.00	34	9.67	1.14	1.62
Spaltungsfähigkeit: 1-2	70	8.11	7.42	39	8.90	7.00	0.56
" 3-5-7	70	18.40	10.68	37	17.46	9.33	0.47
" 4-6-8-9	70	7.51	5.57	39	7.79	4.24	0.30
" A-B	70	7.56	7.14	39	5.59	4.90	1.70
" C-L-P	62	20.08	7.35	34	19.50	8.31	0.34
" D-K-R-T	61	10.49	6.32	33	10.79	5.83	0.23
Cancellation: B-R	51	108.75	23.73	38	112.11	22.18	0.69
" D-K-5	51	68.43	12.49	38	71.37	12.57	1.09
" F-M-Q-X	51	56.78	9.90	38	60.29	11.22	1.54
" C-H-P-T-Z	51	46.41	8.83	38	47.37	8.19	0.53
" 1-5	70	235.11	36.01	36	226.69	34.06	1.19
" 3-7-9	70	182.54	29.56	36	178.33	26.31	0.75
" 2-4-6-8	70	190.24	28.88	36	184.78	28.30	0.94
Incidental memory: right	60	58.25	13.64	29	56.34	14.56	0.59
Incidental memory: wrong	60	37.78	12.65	29	36.45	13.19	0.45

TABLE 7

MEANS, STANDARD DEVIATIONS, AND RELIABILITY OF DIFFERENCES OF PERFORMANCES OF LEPTOSOME AND PYKNIC GROUPS

(*Method of Indices*)

Measure		Leptosome			Pyknic		
	N	Av.	σ	N	Av.	σ	D/σd
Otis	58	46.42	8.66	60	47.53	8.60	0.71
Lecky: score	56	34.60	16.10	57	37.30	20.80	0.77
Lecky: unevenness	57	17.52	5.72	58	17.40	6.52	0.10
Information: score	58	57.48	12.66	60	53.76	10.92	1.74
Information: unevenness	58	26.34	6.33	60	26.34	5.49	0.00
Auditory span	57	8.04	1.14	60	8.18	1.15	0.70
Visual span	47	9.19	1.10	44	9.61	1.31	1.62
Spaltungsfähigkeit: 1-2	51	8.88	6.24	53	7.13	7.14	1.33
" 3-5-7	49	19.35	10.58	53	16.75	10.39	1.25
" 4-6-8-9	50	8.12	5.10	53	7.50	5.39	0.60
" A-B	55	8.25	7.62	53	5.55	5.09	2.18
" C-L-P	45	21.96	9.85	45	16.62	8.49	2.75
" D-K-R-T	47	10.47	6.08	44	9.82	5.33	0.52
Cancellation: B-R	57	108.30	20.00	55	110.80	17.90	0.70
" D-K-S	57	70.45	13.92	56	71.92	13.92	0.55
" F-M-Q-X	57	57.42	11.58	56	60.30	11.76	1.32
" C-H-P-T-Z	57	46.56	9.76	56	45.16	9.68	0.55
" 1-5	52	233.30	33.50	52	230.40	31.30	0.45
" 3-7-9	52	180.80	31.50	52	183.20	24.30	0.43
" 2-4-6-8	52	190.40	29.50	52	189.80	23.60	0.11
Incidental memory: right	43	58.65	13.53	41	60.46	14.60	0.73
Incidental memory: wrong	43	35.63	13.23	41	34.34	11.79	0.47

TABLE 8

MEANS, STANDARD DEVIATIONS, AND RELIABILITY OF DIFFERENCES OF PERFORMANCES OF LEPTOSOME AND PYKNIC GROUPS

(Method of Indices and Diagnosis)

Measure	Leptosome N	Leptosome Av.	Leptosome σ	Pyknic N	Pyknic Av.	Pyknic σ	$D/\sigma d$
Otis	46	45.50	8.56	30	48.00	8.00	1.91
Lecky: total score	46	35.40	17.80	26	35.80	18.80	0.09
Lecky: unevenness	46	17.84	6.20	26	16.00	5.24	1.36
Information: total score	46	57.24	12.96	30	57.78	9.48	0.21
Information: unevenness	46	25.51	5.70	30	25.90	5.19	0.31
Auditory span	46	8.11	1.00	29	8.13	1.00	0.09
Visual span	40	9.15	1.00	23	9.56	1.00	1.37
Spaltungsfähigkeit: 1-2	44	8.98	6.48	27	8.70	7.55	0.16
" 3-5-7	43	19.77	10.44	27	17.92	9.57	0.33
" 4-6-8-9	44	7.86	5.20	27	7.70	4.24	0.14
" A-B	46	8.52	8.00	27	6.00	5.57	1.58
" C-L-P	38	22.50	10.58	23	17.91	7.48	1.98
" D-K-R-T	58	10.68	6.48	22	10.86	5.00	0.12
Cancellation: B-R	44	107.82	18.52	27	114.96	21.93	1.41
" D-K-S	44	68.52	13.03	27	70.96	12.44	0.79
" F-M-Q-X	44	55.95	9.79	27	60.78	10.29	1.96
" C-H-P-T-Z	44	46.59	8.83	27	47.30	7.81	0.35
" 1-5	39	232.95	33.03	25	227.40	25.61	0.78
" 3-7-9	39	180.69	28.54	25	179.84	20.42	0.14
" 2-4-6-8	39	191.26	27.42	25	186.16	26.80	0.74
Incidental memory: right	36	59.50	14.21	20	58.70	14.66	0.20
Incidental memory: wrong	36	36.11	12.69	20	35.65	11.18	0.14

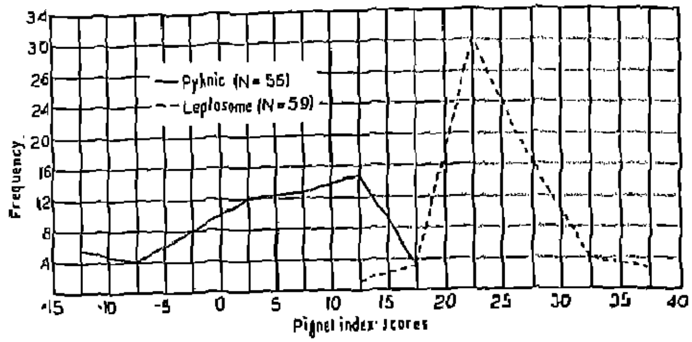


FIGURE 1

DISTRIBUTION OF SCORES OF LEPTOSOME AND PYKNIC GROUPS ON THE PIGNET INDEX (METHOD OF INDICES)

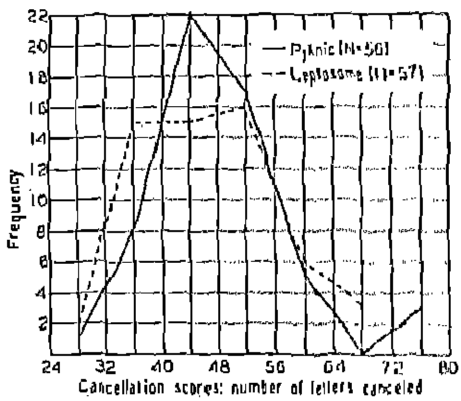


FIGURE 2

COMPARISON OF THE DISTRIBUTION OF CHEST-HEIGHT INDEX IN 79 BARNARD COLLEGE WOMEN, WITH KÜHNEL'S DATA ON 150 "PURE" FEMALE TYPES

tween the physical and mental measures are insignificant. While this procedure of correlating performance scores with *single* physical measures would not be justifiable if a relationship were established between each physical category and mental performance, in the

TABLE 9

CORRELATIONS BETWEEN PHYSICAL INDICES AND PERFORMANCE SCORES (N = 110)

	Pignet index	(Standing Height)[2] / Sitting Height × Weight
Otis	+.01 ±.065	—.1138±.065
Lecky score	+.11 ±.065	+.048 ±.065
Auditory span	—.049±.065	—.091 ±.065
Spalt. 4-6-8-9	+.067±.065	+.087 ±.065
Incidental memory: right	—.110±.065	—.076 ±.061

absence of such evidence it tends to confirm our negative conclusions.

Of similar import are the intercorrelations between the different psychological measures, which the reader will find in Table 10. All of the tests, it will be re-

TABLE 10

INTERCORRELATIONS BETWEEN MENTAL TESTS (N = 110)

	Spalt. 3-5-7	*Spalt. 4-6-8-9*	*Spalt. 1-2*	*Incidental mem.:*
Auditory span	—.203 ±.062	—.284±.059	—.156±.063	
Conc. C-H-P-T-Z	—.1900±.062			—.0624±.064
Incidental mem.	—.0851±.064			
Lecky score		—.097±.064		

called, were designed to measure abilities differentiating the two groups. It should follow that the intercorrelations between the tests, obtained from a population consisting of both types, would be fairly high.

The conclusion is clear, from a survey of the intercorrelations, that the different tasks under consideration have nothing in common with each other. As a matter of fact, we find a surprising consistency of *negative* correlations, although these are small, and significant in only 2 cases out of 7. That the small correla-

tions are not artifacts caused by the low reliability of the separate tests will become clear if the reader will refer to Table 11, giving the reliability of the tests.

TABLE 11

RELIABILITIES OF MENTAL TESTS

Test		Author	Reliability	N.
Otis Self Administering		Otis	.925	125
Lecky Individual: score*		Pallister	.917	209
Visual span*		Anastasi	.7449	225
Cancellation	B-R†	Present	.9412	120
"	D-K-S†	"	.8406	120
"	F-M-Q-X†	"	.8419	120
"	C-H-P-T-Z†	"	.7113	120
Spaltungsfähigkeit Test‡		"	.6227	110

*Reliability was calculated by the split-half method and the application of the Spearman-Brown formula.

†Reliability was determined by correlating the scores of the first and fourth trials with the scores of the second and third trials. The resulting correlation was stepped up by the Spearman-Brown formula.

‡This value was obtained by correlating the scores of the test consisting of the letters C-L-P with the scores of the test including the numbers 3-5-7. The resulting intercorrelation may be accepted as a minimum value of the reliability of each test.

It will be observed that the reliabilities are fairly high. It seems reasonable to conclude, on the basis of the correlations between the tests, that we have evidence for the *absence* of a trait or of a group factor having reference to the ability to analyze groups of impressions. It is such a systematic and powerful central ability that is implied in Kretschmer's hypothesis as well as in his results.

4. *Interpretation of Results.* These findings are in such direct contradiction to those of Kretschmer and other German workers that the question of the comparability of the present experimental situation with

theirs is immediately raised. An answer to this question must be based upon (*a*) the extent to which our tests measure functions similar to those of previous workers in the field, and (*b*) the correspondence between their methods of physical classification and ours.

It goes without saying that we cannot prove the validity of our tests for the measurement of those capacities which, according to Kretschmer, distinguish the leptosome from the pyknic types. It is possible, however, to infer from the structure of the tests the nature of the demands they imposed on the subjects, and the kinds of abilities they probably called into operation. The subsequent sections will be devoted to an analysis of the tests from this point of view.

That there are no differences between the groups on the Otis Self-Administering Test is clearly shown in the results. This finding was expected, in view of previous findings of the same kind [Heidbreder (11), Naccarati and Garrett (27), Garrett and Kellogg (8)]. The result is important, however, for another reason. Most of our tests are measures of intellectual performance of one kind or another. Given this situation, it was quite likely that they might all measure more or less partial aspects of whatever is measured by tests of general intelligence. By correlating the Otis Self-Administering scores with the other scores we were able to show that this situation did not hold. The results are presented in Table 12.

The intercorrelations justify the conclusion that the tests of this study measured abilities independent of

TABLE 12

INTERCORRELATIONS BETWEEN OTIS SELF ADM. AND OTHER TESTS

			r	PE	N
Otis	and	Spaltungsfähigkeit 1-2	—.184	± .044	110
"	"	" 3-5-7	—.270	± .061	108
"	"	" 4-6-8-9	—.315	± .057	110
"	"	Cancellation C-H-P-T-Z	+.137	± .063	110
"	"	Auditory span	+.250	± .058	117
"	"	Incidental memory: right	—.176	± .071	84
"	"	Information score	+.261	± .060	110

those found in tests of "general" intelligence. What abilities did the tests measure?

The Spaltungsfähigkeit tests, or tests of "cleavage" capacity, were modeled upon the tests used by Kibler (14). This investigator exposed tachistoscopically 5 different cards, each containing nonsense syllables varying in color (red, green, blue, and black), and in position. Each card was exposed twice. Before each exposure the subject was directed to attend either to the form or color of the syllables. Kibler reports that the leptosomes were "far superior to the pyknics in keeping distinct the different impressions."

The subjects of this investigation were required to keep track of the number of times each of a set of letters or of numbers was read during a given trial. The test was systematically varied in two directions: (*a*) the number of *different* letters or numbers read within a single trial varied from 2 to 4; and (*b*) the *total* number of letters or numbers included in a single trial varied from 6 to 22. The task confronting the subject was thus to keep together an increasing number of different stimuli as well as a progressively increasing number of total stimuli. It would seem, therefore,

to exemplify what Kretschmer understands by Spaltungsfähigkeit. No reliable differences, however, are present on any part of this test, for any of the methods of classification.

The tests of cancellation were constructed according to the principle already employed with the Spaltungsfähigkeit tests. The number of different letters to be cancelled within any given trial varied from 2 to 5, while the different numbers to be cancelled varied from 2 to 4. Again the test situation required the subject to react to an increasing number of different stimuli within a short period of time. Again the groups are not significantly different on the various parts of the test.

The tests of *auditory span* measure the maximum number of different impressions which an individual can reproduce accurately immediately after presentation. Differences between pyknics and leptosomes in span of perception have been reported by Van der Horst (40), Enke (3), and Kibler (14). These authors exposed tachistoscopically letters arranged in random order, directing the subjects to reproduce as many as possible. The number of letters reported by the subjects are given by Van der Horst and Kibler in the following summary:

	Van der Horst	Kibler
Pyknics	76	42.4
Leptosomes	66	36.8
Circulars	32	32.0
Schizophrenics	28	27.4

Although the *memory span* tests do not measure pre-

cisely the same function as the span of perception, the causes tending to divide the types on the latter test should also operate with the former. The differences, however, were unreliable for our groups.

Two methods of presentation of the memory span material were used. The visual method permitted the simultaneous exposure of all the numbers, while the auditory method necessitated the successive presentation of single numbers. According to the descriptions of the functioning of the two types, we would be led to expect that the pyknic subjects would be benefited by the visual method of presentation, and that the leptosome subjects would find the auditory method more advantageous (at least relatively to the pyknic subjects). From our results it appears that the different methods of presentation produced no reliable difference in the relative standing of the two groups.

The test of incidental memory was designed to measure the relative ability of the two groups to concentrate effectively on a circumscribed task and to resist fairly potent and attractive distractions. According to the investigation of Kibler (14), previously mentioned with reference to the Spaltungsfähigkeit test, the pyknics were better than the leptosomes in reporting those features of the syllables to which their attention was not directed. The differences between the groups are entirely negligible on this test.

The Lecky Individuality Record is a measure of the negative or withdrawal attitude (21). Van der Horst (40) and Kibler (14) claim to have found striking personality differences between the constitu-

tional groups. They studied the self-diagnosis of their subjects, which revealed a strikingly close agreement between pyknic build and cyclothyme disposition on the one hand, and leptosome build with schizothyme disposition. The following summary represents the results of Van der Horst:

	Cyclothyme	Undecided	Schizothyme
Pyknic	94.4%	2.8%	2.8%
Leptosome	12.2%	17.1%	70.0%

From these findings one would predict that leptosome subjects would be more strongly withdrawing or negative. No such difference appears in our results. The other measure with which this test furnishes us is the consistency on the 8 component parts of the test. If, as Kretschmer claims, the tendency of leptosomes to "split" extends into emotional as well as intellectual activities, leading to the frequent formation of complexes among them, we would expect to find some evidence of this in the inconsistency scores. The data fail to confirm the expectation. Results leading to the same conclusions are reported in the second study below.

In view of the descriptions of the pyknic group as relatively mobile in their mental processes and as relatively poorer in concentrating on prescribed tasks, we included a test of general information containing questions concerning 13 different subject matters. The two groups differed negligibly both in total score on the test as well as in their variability on the parts of the test.

In the light of the preceding discussion we feel justified in concluding that our measures approximate as closely as the present state of knowledge permits to those functions which are held by Kretschmer to divide leptosomes and pyknics into two mental types. There seems to be no reason to suppose that the particular tests used by Kretschmer and the other German investigators measured the functions in question more validly than did ours. The tests of this investigation, most of them possessing fair reliability, have in no case indicated a significant difference between the contrasted groups.

The second question of interpretation may be raised with reference to the physical criteria that were employed in the classification of our subjects. This problem has already been discussed in a previous section. It was there pointed out that we made use of a representative sampling of a normal population, that our norms for the constitutional types were in general agreement with those of other workers, that on the basis of a number of indices successfully used in the past, we selected two practically non-overlapping groups, and that the comparison of mental test scores of the two constitutional categories was made using three different methods of classification. We have steered a middle course in the controversy as to whether "general diagnosis" or anthropometric indices are to be preferred by using both, and we have found that the results are exactly the same with either method, or with the two in combination.

It may perhaps be added that one of the writers

spent a month in Kretschmer's laboratory and received training in constitutional diagnosis and in anthropometric measurement; it is a fair assumption that the classification here adopted would not differ from that of Kretschmer to any very marked degree.

5. *Summary of First Study.* (*a*) From a group of 153 subjects, homogeneous in age, sex, educational, social, and economic backgrounds, two sub-groups—leptosome and pyknic—were selected on the basis of physical criteria. A number of mental tests were administered to both groups and their performances were compared.

(*b*) The physical classification of subjects was carried on with three different methods. The results of the three methods were consistent with each other. In addition the norms for the present groups corresponded closely with previously published norms of other investigators.

(*c*) No reliable differences were found between the physical groups on any mental test, using each of the three methods of classification.

(*d*) The physical measures and indices were highly correlated with each other, indicating the possibility of dividing subjects into leptosome and pyknic categories.

(*e*) All the correlations between mental tests were too low to be significant. There is no evidence for any common or central ability running through them.

III

THE SECOND STUDY

A second study, similar in character to the one reported above, was undertaken concurrently, for the purpose of checking the Kretschmer findings in a homogeneous group of women. Barnard College Freshmen and Sophomores were chosen as subjects for the experiment, chiefly because age-differences in the group were negligible and the range of environmental differences presumably small.

A. Procedure

1. *Physical Measurements.* A series of anthropometric measurements had been taken, in September, 1932, by trained workers in the Barnard Physical Education Department for the entire group of 423 students; these were carefully recorded, in February of the following year, from the files of the Department. The Pignet index, the Height-Weight ratio and a modification of Mohr and Gundlach's index (23)

TABLE 13

Distribution of the Height-Weight Ratio in 423 Barnard College Women

Index	Frequency
35-38	6
39-42	17
42-46	48
47-50	95
51-54	123
55-58	77
59-62	39
63-66	14
67-70	2
71-74	2

were then calculated for the entire group and the results distributed. As the preceding table shows, an almost perfect distribution resulted for the Height-Weight ratio; the curves for the other two indices were also nearly identical in shape. (Table 13)

A correlation between the Mohr and Gundlach modification and Height-Weight ratio was further computed and found to be $+.92 \pm .009$; since it was $+.86$ in Mohr and Gundlach's own results, this index was discarded in favor of the simpler ratio.

Invitations to serve as subjects in the experiment were then sent to approximately 175 individuals who fell in the upper and lower ranges of the two normal curves. Students below 16 years of age and above 20 were excluded from the first list. It was hoped, at the inception of the experimental work, that clinical observations similar to those made in the first study might also be made on the Barnard subjects, for the determination of constitutional type. The very limited time at the disposal of the students, however, precluded this possibility, and the investigator was forced to remain content with a mathematical determination of the types, both from the distribution of several indices and from comparison of the data with those of Gottfried Kühnel (20). Seventy-nine students responded to the invitation to be subjects; they were tested over a period of two weeks in April, 1933, in small groups not exceeding twenty a day. The shoulder breadth of each subject was measured before the opening of the actual experiment, since this measurement had not

been included in the original series taken by the Physical Education Department.

2. *The Tests.* Five tests were then administered to the subjects in the order in which they are described below:

Enke's Long Word Test.[2] Two polysyllabic, hybrid words were flashed onto a "daylight" Trans-Lux screen for one-tenth of a second, ten times each. The exposures were mechanically controlled by a camera-shutter device and were flashed from a stereopticon. The words were "extraterritoriality" and "antidisestablishmentarianism." Subjects were given the following directions: "You will now see flashed on the screen letters or words. There will be ten brief exposures, after each of which 15 seconds will be allowed for you to record in the appropriate blank space, *what you see. Please print your answers legibly.*" Scoring was done by classification into "Gestalt," "Analytical," and "Unclassifiable" categories, after the descriptions given by Enke (3), who found schizothymes giving careful, analytical responses, while the cycloids or cyclothymes ran to more synthetic, "Gestalt-qualität" answers.

Spaltungsfähigkeit Test. This test was divided into three parts. (*a*) A series of twelve colored squares was flashed in random order, but rhythmically, onto the Trans-Lux screen, for an entire exposure time of one minute. The squares were made on prepared slides, in which colors were obtained by permitting the "Magic-lantern" light to pass through red, yellow, blue, and green cellophane. Directions were as follows: "You will now see flashed on the screen a series of colored squares. When the entire exposure is complete, that is, when I say 'Ready,' you will record in the appropriate spaces on your blank *how many squares there were of each color.*" Score was obtained by deducting one for each error in a category and then subtracting the total number of wrongs from 12, the maximum score. (*b*) A series of twelve colorless forms was next flashed on the screen, with similar directions and for the same total exposure time. The forms were: square, triangle, oblong, and circle. Scoring was done

[2]In both this and the following tests, the room was only very slightly darkened to avoid any possible error from pupillary adaptation.

in the same way. (*c*) Three separate series of twelve colored forms each were flashed on the screen again in random order and with a total exposure time of one minute for each series. Directions were the same, except that the subjects were requested to record how many "objects there were of each form and of each color." Scoring was done, again, according to categories. It was expected, of course, from the Kretschmer schema, that the leptosomes would do well, and the pyknics poorly, in these tasks.

Incidental Memory. This is the "Logical Memory" test described above in the first study.

Self-Diagnostic Questionnaire. This questionnaire was translated as well as possible from Kibler's (14) diagnostic lists. It is quoted verbatim:

> Here are two sets of adjectives and adjectival phrases, each list being supposed to characterize a certain type of individual. Please read both lists carefully, one at a time, and then indicate by a *check in the appropriate box* which group of adjectives best fits your type of personality. It is not essential that every characteristic apply to you; what we want to know is which *total picture* you think most relevant to you. Will you then indicate by *circling* the heading of either list which type of person you would in general *prefer* to be.
>
> List I—Goodnatured; talkative; friendly; sociable; cheerful; comfortable; easy-going; humorous, with a sense of humor; irritable; easily angered, but quickly calm again; fond of mankind; humane; able to adjust to circumstances, realistic; always thinking of realities; fond of pleasures; fond of eating and drinking; respectable and conscientious; narrow, almost a philistine; busy; active; enterprising but also circumspect; competitive; active, a leader in groups; lively, unconstrained, unaffected; in general, popular; natural; merry, jolly; not formally polite; sometimes miserly, then extravagant; unreserved, not taciturn; practical and determined; fond of children; warmly pious; not rigidly dogmatic; not highly principled; prepared to make concessions, to compromise; fond of nature and of social life; fond of realistic books, nature and travelogues.
>
> List II—Unsociable, except on rare occasions; worrysome or else completely indifferent to dangers; enemy of mankind; silent; monosyllabic; in general, repressed; passionate, irritable; sensitive; impulsive; easily put out of humor; pensive, meditative; humorless, taking everything seriously; given to fantastic dreaming or musing; given to idealising; pedantic, conscientious; having a chronic feeling of insufficiency; longing for distant lands; acting always on logical principles; strongly moralistic; stoical, although at the same time inwardly sensitive; aristocratic; self-sufficient; independent; cynical; sceptical or else deeply religious; bitter; stiff; awkward; shy; unbending; hard; purposeful, tenacious; indifferent; egocentric;

sometimes apparently altruistic but in reality little interested; unsympathetic; in general, not fond of children; given to abstractions; intellectual; living much in the future; given to dreaming; fond of reading idealistic, bizarre and fantastic books.

Now please indicate to what *degree* you think the list you have checked describes you. Let 0 = "slightly," + = "moderately," and ++ = "almost perfectly." Indicate by circling one of the following:

0	+	++
slightly	moderately	almost perfectly

Lecky Individuality Record. This personality inventory is described above. The record was filled out by the subjects at home.

Scholastic Aptitude Test. This test is administered to all entering students of Barnard College as part of the requirements for admission, taking the place of the ordinary intelligence test.

B. Results

1. *Physical Results.* Two additional indices, both of which had been favorably mentioned by Kretschmer (17), Kühnel (20), and Wigert (47), as showing a marked differentiation of the three major body builds, were next calculated for the group of 79 subjects. The indices chosen were the Chest-Shoulder and the Chest-Height,[8] which Kühnel cited as differentiating, respectively, pyknics from athletes and leptosomes, and pyknics from leptosomes, with athletes falling in between. Since Kühnel's curves were based on a study of 150 "pure" constitutional types, all normal women who had first been diagnosed clinically, it was felt that considerable knowledge could be gained by

[8]The Chest-Shoulder index is described above, p. 164. The Chest-Height Index is computed as follows:

$$\frac{\text{Chest-Circumference (cms.)}}{\text{Standing Height (cms.)}}$$

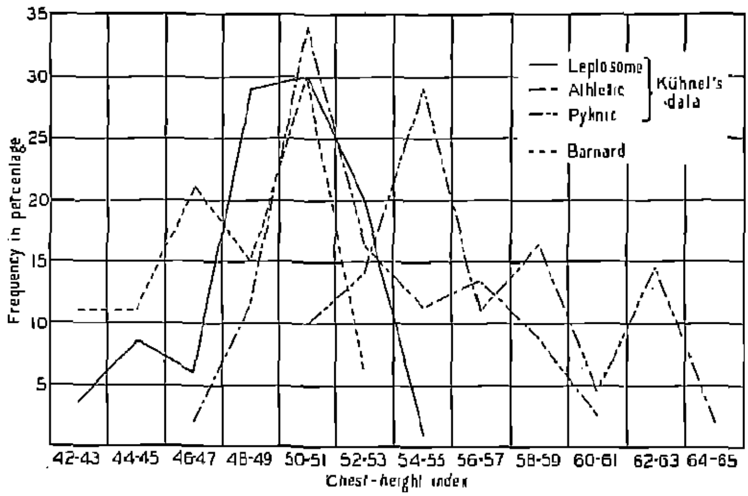


FIGURE 3

DISTRIBUTION OF SCORES OF LEPTOSOME AND PYKNIC GROUPS IN CANCELLATION—CHPTZ (METHOD OF INDICES)

superimposing the curves from the 79 Barnard subjects on Kühnel's graphs.

It appears at once, from a cursory examination of this distribution, that the Barnard curve practically coincides with Kühnel's curve for leptosomes, abutting only slightly into pyknic territory and somewhat more into athletic. All three indices, the Chest-Height, the Pignet, and the Chest-Shoulder, show this result consistently, although in varying degrees.

2. *Correlations between Physical Measures.* The intercorrelations of the indices were then computed, and are in Table 14.

These correlations are all positive in direction and statistically reliable. They show a fair degree of relationship among the indices. It would therefore seem

TABLE 14

INDEX INTERCORRELATIONS

	Chest-Height	Chest-Shoulder	Ht-Wt.
Pignet	+.515±.06	+.488±.06	+.614±.06
Chest-Shoulder	+.557±.06		
Height-Weight	+.742±.06	+.35 ±.07	

safe to conclude that, since the indices intercorrelate positively and show a consistent "skew" when graphed separately on Kühnel's data, there is a definite preponderance of leptoid,[4] and a corresponding absence of pyknoid forms in this Barnard group.

This phenomenon may, of course, be the result of chance, or it may be attributed to some selective factor operating in the group. Kretschmer and Kibler report that leptosomes are prone to shun an experimental situation. The absence of *pyknics* in our group therefore could hardly be laid to direct avoidance of the experiment. Naccarati and Garrett (27), in a study of morphological index in a University group, also found more microsplanchnics than macrosplanchnics. They suggest that since microsplanchny is positively related to intelligence, fewer macrosplanchnics would reach a University group. It is just possible that a similar factor operated in our results, although the zero correlations (to be reported later) between Scholastic Aptitude Test and the four indices make the possibility slim.

In any event, the absence of pyknics among the subjects forced us to abandon our original intention of

[4]The terms "leptoid" and "pyknoid" are somewhat arbitrary designations for the two extremes of the Barnard distribution curve.

treating the various test results in three separate groups —Pyknic, Athletic, and Leptosome—and to substitute correlational method for measuring relationship between temperamental type and body build. The Pearson technique, while not exactly congruent with the Kretschmerian schema, which implies causal interrelationships rather than statistical probabilities, seemed even more advisable since no one index, or combination of them is, in itself, a sufficient criterion of constitutional types. Not only Kretschmer, but Wertheimer and Hesketh (44) have pointed out the difficulty of inductive classification; the "imperfect" intercorrelations of the indices made the difficulty even more apparent. It was felt that whatever differences and relationships existed in the data would probably appear despite a somewhat unsatisfactory technique.

3. *Mental Performance.* The results of the incidental memory and "Spaltungsfähigkeit" tests are presented together in Table 15.

We shall consider first the correlations between physical index and incidental memory. It will be seen at once from the table that the only correlations of possible significance here are those with Height-Weight ratio (—.37) and with Chest-Height index (+.39). The other two are statistically unreliable. The trend of these two former correlations is favorable to Kretschmer's hypothesis, that is to say, those individuals with low scores on the incidental memory test, who evidently attended only to the task in hand, are characterized also, to some extent, by more leptoid body-builds. The correlations are too low, of course, for

TABLE 15

Correlations of Physical Indices with Incidental Memory and Spaltungsfähigkeit

	Incidental* Memory	Colored Squares	Spaltungsfähigkeit† Forms	Colored Forms
Pigner index	—.12 ±.07	—.048±.084	—.084±.054	+.182±.07
Chest-Height index	+.39 ±.07	+.094±.085	+.305±.06	+.27 ±.07
Chest-Shoulder index	—.012±.08	—.258±.07	—.12 ±.07	—.014±.09
Height-Weight ratio	—.37 ±.067	—.140±.07	—.041±.085	—.021±.09

*N=77 †N=79

definite confirmation, and the results are further discredited by the lack of relationship between incidental memory and two such important indices as the Pignet and Chest-Shoulder.

None of the twelve "Spaltungsfähigkeit" correlations is reliable except perhaps that between capacity to "split" a series of forms and Chest-Height index (+.305). The Chest-Height index also shows some relationship to Colored Forms (+.27) although the correlation is not conventionally reliable. The trend of both correlations is this time *unfavorable* to the Kretschmer schema, since a high score in "Spaltungsfähigkeit" should be related to low Chest-Height index (leptoid) rather than to high Chest-Height index (pyknoid).

Lecky Individuality Record. Correlations were computed first between total "Withdrawal" scores and physical index. To be quite certain that results of interest were not neglected through failure to analyse the Record scores according to Lecky's directions, further correlations were then computed for three subgroups of data: Social Adjustment, Cooperation, and Optimism scores. The results are in Table 16.

None of these correlations is statistically reliable. It is interesting to point out in this connection that Pallister (30), working with 209 Barnard College students, also found a correlation of —.059±.05 between Height-Weight ratio and Lecky Record. In our study, the highest correlation with total "Withdrawal" scores, that between the Record and Chest-Shoulder index (+.235), *shows a tendency for high record*

TABLE 16

CORRELATIONS OF PHYSICAL INDICES WITH LECKY TOTAL AND WITH LECKY SUB-SCORES. N=76

	Total Score	Social Adjustment	Sub-Scores: Cooperation	Sub-Scores: Optimism
Pignet Index	—.05 ±.08	—.124±.08	—.14 ±.08	—.10 ±.075
Chest-Height	+.186±.07	+.31 ±.065	+.136±.08	+.17 ±.07
Chest-Shoulder	—.235±.07	—.16 ±.07	—.009±.09	—.21 ±.07
Height-Weight	—.055±.08	—.082±.08	+.009±.09	—.083±.08

scores (negative attitude) to be related to pyknicity. This is, of course, contrary to what one would expect from the Kretschmer schema.

The results of the sub-score correlations appear equally as insignificant as the results from the total Lecky scores. They are, if anything, almost without exception contrary in trend to what might be expected from the Kretschmer schema, since larger "Withdrawal" scores (more "wrong" answers in the three groups) are related to pyknicity rather than to the leptoid habitus.

Scholastic Aptitude Test. No relationship is apparent in these data. The signs of the correlations indicate, however, a contradiction of the results in the literature. Paterson (31), for example, summarizing the work of Naccarati and Garrett, Garrett and Kellogg, Sheldon, Heidbreder, and others, finds a slight but positive correlation between a higher degree of intelligence and microsplanchny. Our results indicate a slight but negative relationship between a higher degree of intelligence and slender body-builds. The small sample in our study and the obvious unreliability of the correlations might, however, easily account for the difference in sign. Two correlations were therefore computed for 409 Freshmen and Sophomores between Scholastic Aptitude score and Pignet index and between score and Height-Weight ratio. These are in Table 17B.

The very slight tendency, therefore, for high Scholastic Aptitude score to be related to the pyknoid body-build is evidently substantiated in the larger series of

TABLE 17A

CORRELATIONS OF PHYSICAL INDICES WITH SCHOLASTIC APTITUDE TEST ($N = 79$)

Pignet index	$-.08 \pm .08$
Chest-Height index	$+.01 \pm .09$
Chest-Shoulder index	$+.067 \pm .08$
Height-Weight ratio	$-.01 \pm .09$

TABLE 17B

CORRELATIONS OF PHYSICAL INDICES WITH SCHOLASTIC APTITUDE TEST ($N = 409$)

Pignet index	$-.19 \pm .033$
Height-Weight ratio	$-.11 \pm .033$

data. The two correlations are not completely reliable, however, and they are so small as to make their discrepancy from those in the literature insignificant as well as unaccountable.

Kibler Questionnaire. The results of this questionnaire and of the Long-Word test to be presented in the next section were treated somewhat more qualitatively. The subjects were divided into two groups—"Leptoid" and "Less Leptoid" or "Pyknoid"—on the basis of two physical indices, the Height-Weight ratio and the Chest-Height index. A comparison was then made between the subjects' self-diagnoses and their actual physical habitus.

It will be seen from Table 18 that there is no tend-

TABLE 18

KIBLER QUESTIONNAIRE

Index	N	Body-Build	Schizothyme	Diagnosis mixed	Cyclothyme
Ht/Wt	11	Pyknoid	3	2	6
	68	Leptoid	12	11	45
Chest-Height	8	Pyknoid	2	2	4
	71	Leptoid	13	11	47

ency for "pyknoid" and "leptoid" individuals to diagnose themselves according to the characteristics they are supposed to possess. On the contrary, there is a decided preponderance of cyclothymic diagnoses in a population the majority of which is leptosomic in body-form! Kretschmer has remarked on the popularity of cyclothymic qualities; it is quite evident that his observation is corroborated, almost too well, in our data.

Enke's Long Word Test. The responses to this test were classified, as reported before, into "Gestalt," "Analytical," and "Unclassifiable" answers. No little difficulty was experienced in making these classifications. It was often felt that categories were being superimposed upon the responses with no other justification than an *a priori* schema. Two rigid criteria were therefore followed throughout the scoring: (*a*) to be called "Gestalt," both series of responses must begin with a complete word, although not necessarily the correct one; (b) to be called "Analytical," both responses must show minute observation for the letters in the exposure, and a gradual "building-up" of the response. All other responses, including those in which an alternation of these "types" of perception was appar-

TABLE 19
LONG WORD TEST

Index	N	Body-Build	Perception Type		Unclassifiable
			Gestalt	Analytical	
Ht/Wt	11	Pyknoid	1	4	6
	68	Leptoid	15	8	45
Chest-Height	8	Pyknoid	1	2	5
	71	Leptoid	15	10	46

ent, were termed "Unclassifiable." The results are in Table 19.

Again it is obvious that body-build bears absolutely no relationship to type of perception. The outstanding facts from our data are the overwhelming preponderance of unclassifiable responses, and the failure of "proper" responses in the extreme portions of the index-curves.

The failure of this entire battery of tests to yield relationships of any significance with indices of body-build prompted the investigators to attempt intercorrelations of the tests themselves, in the hope that the subjects' behavior could at least be shown to "hang together" in the manner outlined by Kretschmer. The results of this series of intercorrelations are in Table 20.

It is at once apparent that the only reliable correlations are those between the "Logical Memory" test and the capacity to "split" forms, and between the "Logical Memory" test and Scholastic Aptitude. The relatively high (for these data) correlation of —.45 between incidental memory and Spaltungsfähigkeit for forms tends to corroborate the Kretschmer hypothesis to the extent that individuals (cyclothymes) who have good incidental memory are also supposed to be unable to split the stream of consciousness. On the other hand, it is difficult to account for the absolute lack of relationship between incidental memory and Spaltungsfähigkeit for colored squares and colored forms. The other reliable correlation, that between Scholastic Aptitude and "Logical Memory" (+.36) is readily under-

TABLE 20

TEST INTERCORRELATIONS

	Colored Squares	Forms	Colored Forms	Lecky	Scholastic Aptitude Test
Logical Memory	—.001±.09	—.45 ±.06	—.005±.09	—.157±.08	+.361±.065
Colored Squares		+.013±.09	+.26 ±.07		+.17 +.08
Forms			+.24 ±.07		+.001±.09
Colored Forms					+.20 ±.08
Lecky					—.038±.09

standable both as to size and direction since incidental memory is probably one of the factors measured in a slight degree by the Scholastic Aptitude Test. The intercorrelations of the Spaltungsfähigkeit tests are interesting from the point of view of the Kretschmer schema. Forms and Colored Forms correlate to the extent of +.24; Colored Squares and Colored Forms to the extent of +.26. Yet Colored Squares and (uncolored) Forms show no relationship at all (+.013 ±.09). Kretschmer believes that, as these correlations would tend to show, sensitivity for color and for form are two separate characteristics of the individual. Further and more elaborate statistical analysis of a larger quantity of data is necessary, however, before definite conclusions on this point may be drawn.

The only other correlation of interest is that between Scholastic Aptitude Test and Lecky Record, the results bearing on the old question of the relationship between intelligence test score and so-called emotional stability. No such relationship is apparent in our data.

As a last check on the validity of these results, a study was made of *average differences* in performance among the three constitutional groups. The 79 subjects were divided into three physical classes on the basis of either one or two of the four indices; the averages of these classes were then simply compared. It was felt that a differentiation might perhaps appear which the somewhat more wholesale correlation technique had obscured. The results of this procedure are in Table 21.

It is clear that these differences are neither statistically reliable nor consistent in trend.

TABLE 21

AVERAGE SCORES

A. Incidental Memory ("Logical Memory")		
Body Build	*Average scores*	*Average scores**
Leptoid	42.08	37.38
Athletoid	42.70	43.64
Pyknoid	42.27	43.38
Index:	Chest-Height	Chest-Shoulder
B. Spaltungsfähigkeit		
	1. Colors—Average scores	*2. Forms—Average scores*
Leptoid	10.59	10.50
Athletoid	11.30	11.30
Pyknoid	11.36	11.25
Index:	Pignet	Chest-Height
	*3. Colored Forms—Average scores***	*Average scores*
Leptoid	12.11	9.72
Athletoid	14.05	9.97
Pyknoid	12.25	10.27
Index:	Chest-Height	Chest-Shoulder
C. Lecky Record		
	Average scores	*Average scores*
Leptoid	32.38	28.04
Athletoid	47.80	34.78
Pyknoid	33.38	33.90
Index:	Chest-Height	Chest-Shoulder
D. Scholastic Aptitude Test		
	*Average scores****	
Leptoid	522	
Athletoid	521	
Pyknoid	554	
Index:		

*$D/\sigma D=1.8$ (leptoid-athletoid).
$D/\sigma D=2.2$ (athletoid-pyknoid).
**$D/\sigma D=2.16$ (leptoid-athletoid).
$D/\sigma D=0.90$ (athletoid-pyknoid).
***$D/\sigma D=1.6$ (athletoid-pyknoid).

4. *Interpretation of Results.* It should by this time be evident that the general conclusion of this second study must be a fairly systematic negation of the Kretschmerian findings we attempted to corroborate.

Not only did a battery of carefully chosen temperamental tests show no relationship to inductively determined physical habitus, but the interrelationships outlined by Kretschmer were absent among the tests themselves. A study of average group differences also yielded negative results.

Three reservations should be made at once, however, to this somewhat sweeping conclusion. First, there was an unexplained absence of pyknics in the group of Barnard subjects. Second, the study suffered further from the fact that no clinical observations of the subjects were made, so that it was necessary to rely on mathematical indices alone for the determination of physical type. Third, correlational techniques were used, for these two reasons, in the statistical treatment of the data. This at once put the entire Kretschmerian concept at a disadvantage so far as *verification* of its tenets is concerned. It is, of course, true that Kretschmer's "intuitive" method of classifying types may be criticized, as well as the statistical invalidity of *his* treatment of results. But since the aim of this particular study was verification of results and verification only, it seems sufficient to point out that differing diagnostic and statistical procedures may have been in part at the basis of the negative results.

It should be noted, however, that the results here reported agree closely with those presented in the first study, in which, it will be remembered, diagnostic and statistical procedures more nearly resembling Kretschmer's were used. The absence of pyknics in the present group produced no difference in the measures of rela-

tionship. It has been pointed out above that the size of a correlation is a function of the number of types included in the population. Yet the correlations for a population containing two types and for a population containing only one are practically the same.

5. *Summary of the Second Study.* (*a*) 79 Barnard College Freshman and Sophomores, each of whom was at either end of a physical distribution curve for 423 fellow-students (using three physical indices) were subjects in the experiment. The group was homogeneous in age, sex, and socio-economic status.

(*b*) The subjects were found to be predominantly leptosome in habitus, according to the norms for women published by Gottfried Kühnel.

(*c*) A series of 5 mental tests, derived from previous investigators, or designed by the authors to reveal cyclothymic and schizothymic characteristics, was administered to the group. The correlations of the test results and 4 physical indices were low, for the most part unreliable, and often contrary in trend to what might be expected from the Kretschmerian schema. An intelligence test administered along with this battery also showed negligible relationship to physical criteria.

(*d*) The intercorrelations of the mental tests were insignificant and inconsistent, showing no "pattern" such as might be expected from the Kretschmer hypothesis.

(*e*) No reliable differences were found between physical groups on any mental tests.

(*f*) The general conclusion of the second study is therefore a negation of the Kretschmer theory.

IV

DISCUSSION OF RESULTS

It is our present purpose to seek for an explanation of the striking discrepancies between the results of Kretschmer and his school and the results of the present investigation. In a preceding section, we have considered two possible causes—differences in physical classification, and inadequacies in the psychological tests—and ruled them out as unlikely. There are, however, a number of important differences in the conditions of experimentation and in the treatment of results between the German type studies and the present. These may have been sufficiently influential, singly or together, to be responsible for the conflicting results. Among the more outstanding of these factors, we may list the following: (*a*) heterogeneity of age; (*b*) heterogeneity of intelligence; (*c*) heterogeneity of social, economic and educational backgrounds; (*d*) heterogeneity of sex; (*e*) absence of measures of dispersion and of the significance of differences between measures; (*f*) unknown test reliabilities. The German studies have, in general, been lax in the control of the above mentioned factors.

Of the various disturbing factors that have been mentioned, that of age is probably the most crucial. That age exerts a modifying influence on physique has been pointed out by many investigators. Kretschmer, indeed, recognizes this, as is evident in the following passage: "When we look at old circulars when they were young, it is particularly remarkable that certain

men and women exhibited quite atypical bodies, longish faces, a narrow build in their twenties, while later on they have developed along distinctly pyknic lines" (18, p. 481). Similarly, Wertheimer and Hesketh (44) found a definite age displacement of body types, the pyknic being more frequent with advancing age. Weidenreich (42) also has shown that individuals of advancing age become more pyknic and less leptosome. Similar conclusions have been reached by Von Rohden and Gründler, Möllenhoff, Kolle and Garvey. It still remains true, of course, that there are different types at each age, and there are some individuals who remain consistent throughout life. The factor of age is, however, sufficiently important to lead to spurious results if not controlled. One condition, then, which studies of type differences must observe is to have the contrasted groups approximately equal in age. The bulk of the German type studies deal with groups grossly heterogeneous in age. Thus Enke (3) in one type study, makes use of a group of 200 subjects, men and women, whose age range was 16-65 years. They were, in addition, recruited from all levels of occupation and intelligence. All the results were treated together, except that they were later analyzed into two age groups. Similar heterogeneity is present in a number of Enke's investigations. Enke and Heising (6), in a constitutional study of normals, employed 240 subjects, male and female, ranging in age from 16-60 years. Two age-groups were compared. It is obvious that this attempt at control is highly inadequate. In another study, Kibler (14) had a group of "middle-

aged" subjects of both sexes, and a normal group of men and women, ranging in age from 20 to 50 years. The instances that have been cited are representative of the kind of controls used in the investigations. In both parts of the present study the requirement of homogeneity with regard to age has been satisfactorily met.

The factor of intelligence has been satisfactorily controlled in the present investigations. In the first study, the groups were equated on the basis of the Otis Self-Administering Examination; it was, furthermore, demonstrated that the tests employed were independent of intelligence test scores. Similarly, in the second study, the scores on the Scholastic Aptitude Test did not correlate appreciably with any other measure, mental or physical. Where, however, there is great heterogeneity of intelligence and where the groups have not been equated for intelligence, we are likely to find both a spurious correlation between intelligence scores and scores on other mental tests and spurious differences between the groups. An actual occurrence of this kind is to be noted in the work of Mohr and Gundlach (23). These authors worked with a group of prisoners, of which the pyknics were grossly poorer on the Army Alpha than the leptosomes. They report a correlation in these groups, between a test of cancellation and the Army Alpha scores of the surprising size of +.795, a result that must be assigned directly to the influence of heterogeneity. While the German studies do not give records of the "intelligence" of their subjects, it is probable, as is apparent from the instances cited in a

previous section, that their groups were exceedingly heterogeneous from this point of view.

Heterogeneity of sex is fraught with danger for all studies of types. On the one hand, differences in bodily development of the sexes make the problems of classification somewhat different for them. On the other hand, possible sex differences in mental test scores should be guarded against. A large number of the German studies included men and women in the same treatment of the results. In the present investigation, the influence of sex was controlled by the simple expedient of making two separate studies and including *only one sex in each.*

Heterogeneity of subjects with regard to social, economic and educational backgrounds may act as disturbing influences in two ways: by distorting the distribution of body types on the one hand, and of test scores on the other. Since differences in social, economic and educational backgrounds are frequently accompanied by differences in intelligence test scores, heterogeneity in these respects will operate as heterogeneity in intelligence. There is, in addition, some evidence of a relation between social class and physique, with a tendency for a greater proportion of leptosomes to be found among the upper social classes. This has been variously explained as due either to earlier maturation in upper classes [Stockard (37)], or to difference in diet and modes of living. [See Weidenreich (42)].

The groups that provided the results of the present investigations possessed a very homogeneous educa-

tional background, and were mostly drawn from the middle classes. No consistent effort has been made by the German workers to control the heterogeneity of these factors. Many of them, as a matter of fact, seem to approximate the maximum amount of heterogeneity possible.

On an equal level of importance with the necessity for the experimental control of the homogeneity of the subjects is to be placed the necessity for determining the dispersion and reliability of the differences of the measures. While the differences reported by German investigators are often large, their failure to give measures of dispersion makes it impossible to evaluate their significance. Strictly speaking, therefore, we cannot say that the *results* of the German investigators are in disagreement with ours, although it is not probable that they agree. It should also be noted that given few cases and many measurements, differences that are "statistically" significant will almost certainly be found as the result of the operation of chance.

All the mental tests of this investigation were administered in groups of twenty or less; most of the German studies were conducted individually. Disadvantages incident to group administration should reflect themselves in the reliability of the tests. These, as given in Table 11, are however fairly satisfactory. The German investigators have not reported the reliability of their tests and it is consequently impossible to come to any definite conclusion concerning the influence exercised by the mode of presentation.

The present study furnishes negative evidence con-

cerning the presence of two psychological types in the normal population. This conclusion must, of course, be interpreted while keeping in mind the limitations of this attempt, particularly with reference to the number of subjects and the range of mental functions. It is possible that the proposed types differ in their reactions to situations with which we did not deal. Obviously it is impossible to overthrow a theory of the scope of Kretschmer's with a single investigation. While it would be premature, at present, to regard Kretschmer's' types as another ballet of ghostly categories, the results of the present investigations, with the use of methods more careful than those of Kretschmer's group, have consistently failed to provide evidence for the presence of types which they claim to have found. The typology may still be important in cases of marked mental maladjustment, but we have not been able to find in it any psychological significance in the case of the normal biotypes.

References

1. Anastasi, A. A group factor in immediate memory. *Arch. Psychol.,* 1930, No. 120. Pp. 61.
2. Enke, W. Die Konstitutionstypen im Rohrschachsen Experiment. *Zsch. f. d. ges. Neur. u. Psychiat.,* 1927, 108, 645-674.
3. ———. Experimentalpsychologische Studien zur Konstitutionsforschung. *Zsch. f. d. ges. Neur. u. Psychiat.,* 1928, 114, 770-794.
4. ———. Die Psychomotorik der Konstitutionstypen. Leipzig: Barth, 1930. Pp. 56.
5. ———. Die Affektivität der Konstitutionstypen im Psychogalvanischen Versuch. *Zsch. f. d. ges. Neur. u. Psychiat.,* 1932, 138, 231-237.
6. Enke, W., & Heising, L. Experimenteller Beitrag zur Psychologie der "Aufmerksamkeitsspaltung" bei den Konstitutionstypen. *Zsch. f. d. ges. Neur. u. Psychiat.,* 1928-29, 118, 634-644.
7. Farr, C. B. Bodily structure, personality, and reaction types. *Amer. J. Psychiat.,* 1927-28, 7, 231-234.
8. Garrett, H. E., & Kellogg, W. N. The relation of physical constitution to general intelligence, social intelligence, and emotional instability. *J. Exper. Psychol.,* 1928, 11, 113-129.
9. Garvey, C. R. Comparative body build of manic depressives and schizophrenic patients. *Psychol. Bull.,* 1933, 30, 567-568.
10. Haarer, —. Inaugural-Dissertation, Tübingen (see Kretschmer, 17).
11. Heidbreder, E. Intelligence and the height-weight ratio. *J. Appl. Psychol.,* 1926, 10, 52-62.
12. Jislin, S. G. Körperbau, Motorik und Handschrift. *Zsch. f. d. ges. Neur. u. Psychiat.,* 1925, 98, 518-523.
13. Kelley, T. L. Crossroads in the mind of man. Stanford University, Calif.: Stanford Univ. Press, 1928. Pp. vii+238.
14. Kibler, M. Experimentalpsychologische Beitrag zur Typenforschung. *Zsch. f. d. ges. Neur. u. Psychiat.,* 1925, 98, 524-544.

15. Kolle, K. Körperbau der Schizophrenen. *Arch. f. Psychiat. u. Nervenkh.*, 1924, 72, 40-88; 1925, 75, 21-61.

16. ———. Klinische Beiträge zum Konstitutionsproblem. *Arch. f. Psychiat. u. Nervenkh.*, 1926, 77, 183-238.

17. Kretschmer, E. Körperbau und Charakter. (10th ed.) Berlin: Springer, 1931. Pp. 240.

18. ———. Physique and character. (Trans. by J. H. Sprott.) New York: Harcourt, Brace, 1925. Pp. xiv+268.

19. Kroh, O., Bayer, G., Dambach, K., Lutz, A., & Vollmer, O. Experimenteller Beiträge zur Typenkunde. *Zsch. f. Psychol.*, 1929, 14.

20. Kühnel, G. Die Indexberechnung der Weiblichen Körperbautypen. *Zsch. f. d. ges. Neur. u. Psychiat.*, 1931, 134, 528-555.

21. Lecky, P. The negative attitude. Unpublished Ph.D. dissertation, Columbia University, 1932.

22. Martin, R. Lehrbuch der Anthropologie. Jena: Fischer, 1914. Pp. xvi+1181.

23. Mohr, G. F., & Gundlach, R. The relation between physique and performance. *J. Exper. Psychol.*, 1927, 10, 117-157.

24. Möllenhoff, F. Zur Frage der Beziehungen zwischen Körperbau und Psychose. *Arch. f. Psychiat. u. Nervenkh.*, 1924, 71, 98-127.

25. Munz, E. Die Reaktion des Pyknikers im Rorschachsen Psychodiagnostischen Versuch. *Zsch. f. d. ges. Neur. u. Psychiat.*, 1924, 91, 26-92.

26. Murphy, G. Types of word association in dementia praecox, manic depressives, and normal persons. *Amer. J. Psychiat.*, 1923, 3, 539-571.

27. Naccarati, S., & Garrett, H. E. The influence of constitutional factors on behavior. *J. Exper. Psychol.*, 1923, 6, 455-465.

28. Oseretsky, N. Körperbau, Sanitäre Konstitution und Motorik. *Zsch. f. d. ges. Neur. u. Psychiat.*, 1926, 106, 120-135.

29. Ossipowa, E. A., & Ssucharewa, G. E. Körperbau, Motorik, und Charakter der Oligophrenen. I. Knaben; II. Mädchen. *Zsch. f. d. ges. Neur. u. Psychiat.*, 1928, 114, 1-37.

30. PALLISTER, H. The negative or withdrawal attitude. *Arch. Psychol.*, 1933, No. 151. Pp. 56.

31. PATERSON, D. G. Physique and intellect. New York: Century, 1930. Pp. 304.

32. RORSCHACH, H. Psychodiagnostik. (2nd ed.) Berne: Huber, 1932. Pp. 230.

33. SALLER, K. Konstitution und Rasse beim Menschen. *Zsch. f. d. ges. Anat.*, 1929, 3 Abt., 28, 250-422.

34. SCHNECK, M. R. The measurement of verbal and numerical abilities. *Arch. Psychol.*, 1929, No. 107. Pp. 49.

35. SCHOLL, R. Untersuchungen über die Teilinhaltliche Beachtung von Farbe und Form bei Erwachsenen und Kindern. *Zsch. f. Psychol. u. Physiol. d. Sinneror.*, 1927, 101, 225-280.

36. SMITH, G. M. Group factors in mental tests similar in material or in structure. *Arch. Psychol.*, 1933, No. 156. Pp. 56.

37. STOCKARD, C. R. The physical basis of personality. New York: Norton, 1931. Pp. 320.

38. THORNDIKE, E. L. Educational psychology. Vol. II. New York: Teach. Coll., Columbia Univ., 1914. Pp. x+408.

39. TIMMER, A. P. Die Schizothymen und Cyclothymen Temperamente Kretschmers im Lichte der Pawlowischen Bedingten Reflexe, Betrachtet. *Zsch. f. d. ges. Neur. u. Psychiat.*, 1931, 133, 341-380.

40. VAN DER HORST, L. Experimentell-psychologische Untersuchungen zu Kretschmers "Körperbau und Charakter." *Zsch. f. d. ges. Neur. u. Psychiat.*, 1924, 93, 341-380.

41. VON ROHDEN, F., & GRÜNDLER, W. Über Körperbau und Psychose. *Zsch. f. d. ges. Neur. u. Psychiat.*, 1925, 95, 37-78.

42. WEIDENREICH, F. Rasse und Körperbau. Berlin: Springer, 1927. Pp. 187.

43. WERTHAM, F. Experimental type psychology. *Arch. Neur. & Psychiat.*, 1930, 24, 605-611.

44. WERTHEIMER, F. I., & HESKETH, F. The significance of the physical constitution in mental diseases. London: Baillière, 1927.

45. WESTPHAL, K. The use of indices as an auxiliary method in the establishment of physical types. *Human Biol.*, 1931, 3, 420-429.

46. WIERSMA, E. D. Körperbau Verschiedener Rassen und Konstitutionen. *Zsch. f. angew. Psychol.*, 1929, 33, 136-184.
47. WIGERT, V. Versuche zur Anthropometrischen Bestimmung der Körperbautypen. *Zsch. f. d. ges. Neur. u. Psychiat.*, 1933, 143, 651-700.

UNE ÉTUDE EXPÉRIMENTALE DES TYPES CONSTITUTIONNELS

(Résumé)

Il s'agit d'essayer d'employer des méthodes expérimentales dans l'étude des différences psychologiques entre les types constitutionnels. Le point de départ est la théorie de Kretschmer, qui croit qu'il existe une relation intime entre le physique et la mentalité. Ses croyances ont reçu quelque confirmation, non seulement dans la clinique psychiatrique, mais aussi dans le laboratoire psychologique, et les hommes de science allemands ont fait beaucoup d'étude intéressantes dans ce domaine. Cette investigation-ci se sert de quelques-unes de leurs techniques, ainsi que de quelques nouvelles, faites pour mesurer les capacités qui auraient selon Kretschmer une base constitutionnelle. On a choisi des types picniques et leptosomes d'entre les étudiants de Brooklyn College et de Barnard College et l'on a fait le diagnostique et selon l'observation qualitative et par l'usage des indices anthropométriques. Les groupes dans les deux collèges ont été égalisés pour l'âge, l'état social, l'état économique et l'intelligence. Chaque sujet a subi une batterie de tests y compris des mesures de la portée de la mémoire, de la mémoire incidentelle, de la "Spaltungsfehigkeit," ou capacité de se fendre, et un inventaire personnel, et on a analysé les données pour les différences constitutionnelles de type. Les résultats ont été entièrement négatifs, car nulle différence psychologique signifiante ne s'est montrée dans l'usage d'aucune de ces techniques expérimentales. La divergence entre ces résultats et ceux des expérimentateurs allemands est difficile à expliquer, mais il est possible que certaines différences des processus expérimentaux et statistiques, ainsi que dans l'égalisation des groupes, en puissent être responsables. En tout cas, tant qu'elle va, cette étude combat la possibilité d'arriver à des différences constitutionnelles fondamentales de type selon les méthodes de la psychologie expérimentale.

KLINEBERG, ASCH ET BLOCK

EINE EXPERIMENTELLE UNTERSUCHUNG DER KONSTITUTIONSTYPEN

(Referat)

Der Versuch wird gemacht, experimentelle Methoden zur Untersuchung psychologischer Differenzen zwischen den Konstitutionstypen anzuwenden. Der Ausgangspunkt ist die Lehre von Kretschmer, der glaubt, dass es eine enge Beziehung zwischen dem Körperbau und dem Geist gibt. Seine Ansichten haben einige Bestätigung nicht nur in der psychiatrischen Klinik, sondern auch in dem psychologischen Laboratorium gefunden, und deutsche Forscher haben interessante Untersuchungen auf diesem Gebiet geleistet. Die vorliegende Untersuchung bedient sich eines Teiles ihrer Methodik sowie einiger neuer Methoden, welch zum Messen Fähigkeiten aufgestellt wurden, die Kretschmer als konstitutionell begründet ansah. Pyknische und leptosome Typen wurden unter Studenten der Brooklyn und Barnard Hochschulen ausgewählt, indem die konstitutionelle Diagnose sowohl durch Indizes gemacht Wurde. Die Gruppen aus beiden Universitäten Wurden qualitative Beobachtung wie durch die Verwendung anthropometrischer hinsichtlich des Alters, der sozialen und wirtschaftlichen Lage und Intelligenz gleichgemacht. Jeder Versuchsperson wurde eine Batterie von Tests gegeben, die Masse des Gedächtnisumfangs, nebensächlichen Gedächtnisses, der Spaltungsfähigkeit, und ein Persönlichkeitsinventar umfasste; und die

Daten wurden für konstitutionelle Typenunterschiede untersucht. Die Ergebnisse waren ganz negativ, keine bedeutenden Differenzen erschienen durch die Anwendung dieser experimentellen Methoden. Der Widerspruch zwischen diesen Resultaten und denen der deutschen Forscher ist schwer zu erklären, aber es ist möglich, dass gewisse Differenzen in dem experimentellen und statistischen Verfahren sowie die Ausgleichung der Gruppen dafür verantwortlich sind. Jedenfalls nimmt die vorliegende Arbeit, soweit sie geht, Stellung gegen die Möglichkeit der Erlangung grundlegender Typenunterschiede durch die Methoden der experimentellen Psychologie.

KLINEBERG, ASCH UND BLOCK

$7.00 per volume
Single numbers $2.00

MONTHLY
Two volumes per year

October, 1934
Volume XVI, No. 4

GENETIC PSYCHOLOGY MONOGRAPHS

Child Behavior, Animal Behavior, and Comparative Psychology

EDITED BY
CARL MURCHISON

Worcester, Massachusetts

Entered as second-class matter December 1, 1925, at the post-office at Worcester, Mass., under Act of March 3, 1879

Journal of General Psychology

EDITED BY

CARL MURCHISON, *Clark University*

Quarterly, published in January, April, July, and October. Devoted primarily to experimental, theoretical, clinical, and historical psychology. Manuscripts may be sent to any member of the Editorial Board, or may be sent directly to the general editorial office. All subscriptions and business communications should be sent directly to the Clark University Press. Beginning January, 1932, one thousand pages (two volumes) annually. Per annum $14.00; per volume $7.00; single numbers $4.00. Complete sets from 1927 at $[illegible] and $[illegible] per volume, plus transportation.

The Pedagogical Seminary and

Journal of Genetic Psychology

EDITED BY

CARL MURCHISON, *Clark University*

Quarterly, published in March, June, September, and December. Devoted to child behavior, animal behavior, and comparative psychology. Manuscripts may be sent to any member of the Editorial Board, or may be sent directly to the general editorial office. All subscriptions and business communications should be sent directly to the Clark University Press. Beginning January, 1931, one thousand pages (two volumes) annually. Per annum $14.00; per volume $7.00; single numbers $4.00. Complete sets from 1891 at $[illegible], $[illegible], and $[illegible] per volume, plus transportation.

CLARK UNIVERSITY PRESS
Worcester, Massachusetts

$7.00 per volume
Single numbers $2.00

MONTHLY
Two volumes per year

October, 1934
Volume XVI, No. 4

GENETIC PSYCHOLOGY MONOGRAPHS

Child Behavior, Animal Behavior, and Comparative Psychology

THE DEVELOPMENT OF A BATTERY OF OBJECTIVE GROUP TESTS OF MANUAL LATERALITY, WITH THE RESULTS OF THEIR APPLICATION TO 1300 CHILDREN*

From the Institute of School Experimentation, Teachers College, Columbia University

By

WALTER N. DUROST

*Accepted for publication by Leta S. Hollingworth of the Editorial Board and received in the Editorial Office, October 17, 1934.

Worcester, Massachusetts

Entered as second-class matter December 1, 1925, at the post-office at Worcester, Mass., under Act of March 3, 1879.

ACKNOWLEDGMENTS

The author gratefully acknowledges his indebtedness to the following persons who have made this study possible: To Dr. Otis W. Caldwell, Director of the Institute of School Experimentation, for the appointment to the Institute to carry out this investigation and for his encouragement and friendly criticism throughout the period of our association; to Dr. Helen M. Walker of Teachers College, Columbia University, for invaluable assistance in dealing with the statistical aspects of the study; to Dr. Leta S. Hollingworth, also of Teachers College, from whom the suggestion for the study originally came, for helpful advice while it has been in progress; to Dr. Jack W. Dunlap of Fordham University, who read a portion of the manuscript and facilitated the work by constructing the nomograph described in the appendix; and finally to the superintendents and principals without whose help this study would have been impossible, Dr. John K. Bolan and Mr. Emanuel F. Van Dam, district superintendents of the New York City schools, Mr. Leon J. Deming, superintendent of the Oyster Bay public schools, Miss Jane E. Monohan, formerly principal of P.S. No. 42, Brooklyn, Mr. August Lodato, formerly principal of P. S. No. 44, Brooklyn, and Mr. Samuel Viertel, principal of P. S. No. 43, Manhattan.

WALTER N. DUROST

INSTITUTE OF SCHOOL EXPERIMENTATION
TEACHERS COLLEGE, COLUMBIA UNIVERSITY
NEW YORK CITY

CONTENTS

I

THE MAJOR PROBLEMS FACING INVESTIGATORS OF MANUAL LATERALITY

INTRODUCTION

The widespread interest in the problem of leftness, whether of hand, eye, or of the body as a whole, is indicated in the surprising number of articles which have been published on the subject in recent years. *The Psychological Bulletin* for February, 1933, presents a bibliography of 219 articles under the head "Laterality of Function," happily relieving the present writer of the task of extended bibliographical discussion. It does not, however, relieve him of the necessity of stating his viewpoint on some of the many aspects of this general problem which have direct or indirect bearing on the work herein presented.

HANDEDNESS DEFINED

If one were to ask a layman to define handedness, the invariable answer would be that a left-handed person is one who uses his left hand to do most of the things which involve a choice of hands, with a converse definition for the right-handed. This definition of handedness might be called the preferential definition and has been employed in many, probably a majority of the investigations which have been made to date. There have been, however, some very important deviations from this "preferential" definition which must be noted. One such important deviation holds that handedness is an expression of a fundamental

asymmetry of all the bilateral bodily structures or functions. As such, handedness would be an inherited characteristic which ought to follow definite laws of descent, probably as a Mendelian trait. The preferential use of one hand is, however, so much a matter of training and of environmental pressures that any measures of preferential handedness would not be likely to yield results following known laws of inheritance. Therefore those who stress the importance of the determining of the inherent qualities of handedness unaffected by environment have sought more or less indirect methods of measurement from which native handedness might be deduced. The chief method of indirect measurement has been the determination of eye dominance on the assumption that eye and hand would agree if nature were undisturbed. The devices used have ranged in complexity from simple sighting tubes to elaborate laboratory apparatus. Other indirect methods have made use of the precedence of action or nerve currents. An assumption underlying all indirect methods measuring handedness is that laterality, in the natural state, is consistent for all bilateral functions. This is a dangerous assumption to make since it can not be made subject to experimental verification. Furthermore, the assumption that eye dominance or the precedence of action currents are free from environmental influence is not beyond challenge, although the amount of influence is probably at a minimum in such cases. Some valuable evidence has been produced by these methods, but the limitations of the laboratory so reduce the number of cases studied

that the process of gathering data is a slow and laborious task.

A third approach which is, in a sense, mid-way between these two is the measurement of qualitative and quantitative differences between the hands (or arms or sides) by means of objective tests of skill, muscular coordination, or speed. Strength of grip, speed in tapping, accuracy in throwing, steadiness in guiding a stylus are examples of such tests. The difference between the two bilateral organs being compared can be generalized into some kind of standard score in the form of a ratio. Thus expression is given to a superiority that actually resides in one of the organs itself in the way of better coordination, better muscle tone, stronger muscles, or a combination of all these things. This type of test obviously is not free of the environmental influences which those who use some one of the indirect means of testing are trying so hard to escape, but it does have other very real advantages which will be discussed at greater length later in the monograph. Handedness defined from this point of view, then, would involve the demonstration of actual superiority of one hand over the other in some physical attribute or skill.

Three definitions of handedness have been presented: (1) handedness consists in the preferred use of one hand in habitual performances such as writing, throwing, eating, etc; (2) handedness is an innate disposition based upon structural or neurological factors determined before birth, definitely inheritable, probably as a Mendelian trait, and can therefore be measured only

by some means which will heavily discount the effect of environment and training; and (3) handedness expresses itself as a functional or structural, or both functional and structural, superiority of one organ over its complementary organ as measured by some objective test. The use of any one or all three of these definitions is justifiable so long as the writer informs his readers what definition he has in mind. This study is chiefly concerned with the first and last of these definitions.

Environment vs. Heredity

The problem of environment *vs.* heredity is too important to pass over without some further discussion on it, although it is impossible to go into the matter as fully as the importance of the subject would justify. The problem is not essentially different in nature from that which has raged so furiously in the field of intelligence testing for some years. Furthermore, the end result promises to be the same, namely a compromise between opposing forces, neither of which can prove its point.

The evidence on handedness as an inherited characteristic comes from several sources. There is the observation of very young children such as goes on in Gesell's Clinic at Yale (5). Gesell finds that handedness, either left or right, is definitely established at the end of the second half of the first year. On the other hand M. C. Jones (8) found no preferential choice in infants from 80 to 300 days of age tested for thumb opposition and reaching. Wellman (12) and Hein-

lein (7) report increase in the superiority of the dominant hand as preschool children grew older. However, evidence from this source can never be conclusive for two reasons: (1) the influence of environment begins at the moment of conception, and (2) handedness, calling for a considerable specialization of nerve cells, may be a matter of delayed maturation.

A second source of information is the frequency of appearance of left-handedness among siblings, parents, grandparents, and other relatives compared with the generality. Only one or two studies have been made along this line due to the excessive difficulty in obtaining the necessary data. It is almost impossible to go back of the grandparents of children now in school. Furthermore, it is difficult to assay the effect of the example of left-handed parents and older brothers and sisters on younger children. Nevertheless, comparative studies of relatives, especially large living family groups, can contribute much more than they have to the evidence on inheritance of handedness. The most valuable study along this line is that of Chamberlain (2) who questioned a population of 12068. He found a significantly greater percentage of left-handedness in families in which one or both of the parents were left-handed than in families in which both parents were right-handed.

A third source of evidence is the study of identical twins. Other biological problems enter to make the evidence at this point unconvincing, although this would seem to be a promising line of approach.

A fourth kind of evidence seems to the writer to offer

the most effective attack on the problem. This is the study of handedness in animals. The handedness of rats has been shown by several studies and the author has demonstrated to his own satisfaction that it can be shown in cats, although the experiment needs to be tried again with a larger number of cases. If rats of predetermined handedness could be bred through several generations with control and experimental groups, very valuable evidence might be contributed on this point.

When all the evidence is examined can one say that the inheritance of handedness or of laterality of function is established? It is the writer's opinion that this can not be said unequivocably but that the weight of the evidence favors inheritance as against purely environmental determination of functional laterality. This is important. The reason for its importance lies in the fact that no matter how objective and impartial one is in the collection and analysis of data, the interpretation of results is always from the point of view of one's personal bias. The writer believes that the occasional advantages of being left-handed are so far outweighed by the obvious disadvantages that he would recommend the exertion of environmental pressures early in the child's life to enforce right-handedness *if he believed environment were the major causative* force at work. However, the fact that the weight of the experimental evidence favors the theory of inheritance, in the writer's opinion, plus the fact that leftness has been known to persist in many cases against the most strenuous efforts to dislodge it, leads the writer to

recommend, rather, that nature be allowed to take her course and once that course has become clearly established, that environmental pressures should be made to reenforce it rather than to oppose it.

EDUCATIONAL IMPLICATIONS

The justification of any further discussion of the educational implications of left-handedness in a monograph otherwise devoted to the presentation and analysis of objective data lies only in the fact that in the writer's mind the reasons for our concern with this problem, other than a purely academic interest in an interesting research, have not been sufficiently clear in the minds of many workers in this field. Somewhere in the neighborhood of five in every one hundred school children are found to be definitely left-handed by practically all investigators. Even if we ignore for the moment the conclusions of some important investigations to the effect that left-handedness is somehow linked with reading difficulty, with stuttering and stammering, and with lack of physical coordination, there are still the obvious handicaps in the way of physical equipment designed for right-handed children, incorrect lighting for writing and reading, and occasional enforced emphasis on the right hand in skills other than writing that make it important to assay the handedness of children on wide enough scale to enable us to generalize concerning the success with which the left-handed child, not obviously maladjusted, is meeting these handicaps. And this, among other things, is the purpose of this research.

The Purpose and Scope of This Research

This study makes no use of the indirect methods of measuring handedness. Handedness is measured chiefly by means of a battery of objective tests, the development of which constitutes one major phase of this study. A second major phase is the presentation of distribution curves of handedness for each of the tests of the battery and for the composite score, which is a combination of equally weighted scores obtained with the separate tests. A third phase is the development of a short criterion questionnaire which might be used to select a known proportion of the left-handed from another population for further study; and finally handedness is related to other characteristics for which data were available such as school achievement and intelligence.

II.

NATURE OF TESTS USED WITH RESULTS OF PRELIMINARY TRIALS

Peculiar Sampling Problems

The measurement of laterality offers peculiar sampling problems not often met in educational research due to the fact that leftness characterizes only a small proportion of the whole group. In the case of handedness this proportion seems by general agreement to be somewhere in the nature of 5%, while estimates vary in the case of eyedness and other bilateral functions. Thus one must have a total population of sample of one thousand cases to expect a population of fifty left-handed children. The necessity of testing the entire thousand cases can be avoided, of course, by using some arbitrary criterion for the selection of left-handed children such as parent's choice or the hand used in writing; Haefner (6, pp. 8-11) did this. He surveyed roughly a thousand cases in grades four to seven, using as his criterion the children's answers to three questions: "Which hand do you write with?" "Which hand do you throw with?" and "Which hand do you reach with?" On the basis of the answers to these questions plus the teachers' judgment of the handedness of their pupils, he chose an experimental population of 68 children about evenly divided between boys and girls, and matched them with 68 right-handed children of the same age, sex and grade. These two groups constituted his experimental and control groups and all comparisons were made between these two populations.

The chief drawback in this procedure is that the criterion of selection can not help but definitely bias the sample of left-handed children obtained. Thus Haefner's study might more accurately be entitled "The Educational Significance of Left-Handedness Determined by Writing, Throwing, and Reaching, Plus Teachers' Judgment." If the criterion is fallible, that is, if it does not select a truly representative sample of the left-handed population, any conclusion drawn from comparisons so made are faulty.

Distribution Curves Desirable Outcome

Other even more fundamental reasons militated against using a similar procedure in this research. In the first place, it was felt that a highly desirable outcome of the study would be a distribution curve which would represent handedness as a continuous variate ranging from extreme left-handedness to extreme right-handedness.

The distribution curve which describes handedness has been discussed several times in the literature. However, many questions concerning this distribution curve remain to be answered, justifying much additional work along this line.

In the second place, ambidexterity has been dealt with in the studies mentioning it at all as a matter of mixed preferential handedness. That is, the subject does this act with one hand and that act with the opposite hand, etc. If a continuous distribution of handedness based upon objectively scored tests could be established, it would be possible to compare, not only the

extremes of leftness and rightness with each other but also with a group whose test scores indicated no choice of handedness. If such a distribution of scores could be related to other data concerning change of handedness very important information might be discovered concerning what, in this writer's opinion, is the most critical section of the distribution.

Finally, if a large enough group previously unselected as regards handedness could be studied with a battery of objective tests, it might be possible to establish some criterion which could be expected to select the left-handed and ambidextrous cases from another unselected population for further study. Accordingly, an effort was made to devise a battery of tests and to administer them to a minimum of one thousand cases over a range of several grades of the elementary school, for the following reasons:

1. To establish the distribution curve of handedness with a series of objective tests.
2. To establish with greater objectivity the significance of ambidexterity.
3. To establish norms of handedness together with a criterion which would reliably select cases falling in the three critical ranges of the distribution, i.e., the extremes of leftness and rightness and the ambidextrous.

DESIRABLE CHARACTERISTICS OF HANDEDNESS TESTS

Such a test battery should meet the following requirements:

1. It must be made up of group tests for the sake of speed in accumulating data.
2. It must be of known reliability.
3. It must be a valid measure of handedness not in the sense of a high correlation with a criterion so much as in its being obviously a test of skill not greatly favoring either hand, but permitting of a fair comparison of the relative achievement of the hands.
4. It must be simple to administer so that persons with relatively little training could administer the tests.
5. It must not take so long to administer that it would seriously interfere with the daily routine of the school, or be greatly affected by the fatigue factor.
6. It must be as objective to score as a good intelligence test.
7. It should be interesting to those taking it in order to secure good motivation.

No claim is made here that the battery of tests finally used meets all these criteria with equal adequacy. As will be seen it obviously does not, but it does meet them with sufficient exactness to justify the use to which it was put.

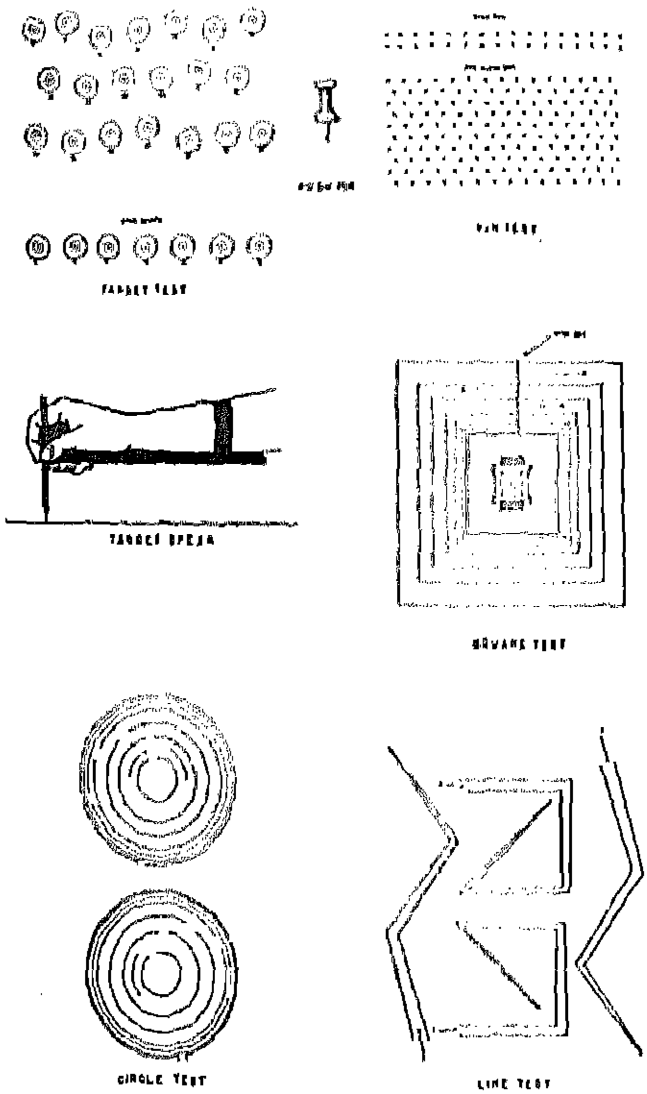
Nature of Tests Finally Selected

In surveying the literature on handedness tests, it was felt that certain of the tests commonly used in the psy-

chological laboratory offered the most promising beginning point. They had been used as handedness tests before, though not primarily so, and they were simple and objective. The greatest drawback was that they were individual tests and as such would be impossible to give to the large experimental group being considered. The tests in mind were the tapping test, using a metal stylus and plate in series with a counter, the target board test, and a steadiness test, in which a metal stylus is inserted between two converging metal strips until an electrical contact is made.

With this as a starting point an effort was made to devise tests along similar lines which would meet the criteria stated above. After considerable experimentation a battery of four tests was constructed and tried out in the elementary school of Oyster Bay, Long Island. Later a fifth test was added and tried out in the same school with the same classes. Figure 1 will give an idea of what these tests were like. All preliminary forms of the tests were mimeographed on 24 lb. mimeograph paper. The first test was called the Pin Test because it made use of a large Moore pushpin. The test proper consisted of ten lines of the small letter *o*, 15 to a line. The lines were staggered to break up any tendency to rhythmic movements. The tests were pasted upon double-thick corrugated board, with one test on each side of the board, to give sufficient thickness to absorb the blow of the pin and protect the desk surfaces. The object of the test was to puncture as many of the small circles as possible in the time allowed, which in this case was one minute. The direc-

FIGURE I



HANDEDNESS TESTS — PRELIMINARY FORMS

tions were to "make a hole in each circle with the pin. Puncture as many circles as possible in the time allowed." The test was done first with the writing hand and then with the non-writing hand. The test was intended to be roughly analogous to the tapping test of the psychological laboratory. It was supposed to measure the efficiency of the small muscle movements of the hand in terms of both speed and accuracy.

The second test was called the Target Test, and the test proper consisted of twenty small, irregularly spaced targets as may be seen in the illustration. This was also mounted on two thicknesses of corrugated cardboard. The target spear, so called, was the implement used in this test. It consisted of a shaft, approximately 10 inches long and ½ inch square, through which a ¼-inch dowel was mounted at right angles about a quarter of an inch from one end. Approximately two inches from the other end was fastened a loop of elastic which could be adjusted in size. In operation the arm was slipped through the elastic loop until the thumb came to the end of the shaft with the fingers on top, and the elastic loop was made firm around the upper forearm. This contrivance had the effect of locking the wrist so that the movement made in striking at the targets was a movement of the forearm rather than the hand and the larger muscles were called into play. Three shots were allowed at each target as the examiner counted 1-2-3. The first test was done with the writing hand and the second with the non-writing hand so that any transfer-practice effect there might be would accumulate to the advantage of the non-writing hand.

This test was intended to be roughly analogous to the target test of the psychological laboratory. (The bull's eye of a target counted three, the next circle two, and the outside circle counted one. The final score was the sum of the scores on the separate targets.)

The remaining tests of the original battery were all variations of a common pattern. They were designed to measure steadiness in the same way that the slot test measures it in the laboratory. A pencil was used in place of a stylus and the line drawn marked the path followed between the lines or "walls" of the test.

The Line Test was the first one of this type to be devised and it followed the thought of the slot test literally. It consisted of a series of paired converging lines forming various patterns. There were four patterns or units in the test so devised that if the paper were kept in the same position all the time, the person taking the test would have to draw the lines in a variety of directions. The test was scored in two ways: the distance traveled in centimeters to the point where the pencil line first touched the side walls or printed lines of the test, and the number of such contacts made from the point of first contact to the end of the test. This test proved to have unsuspected difficulties in administration. It was almost impossible to keep the children from moving the paper so that they were always drawing the line in one position, usually horizontally. Furthermore, the figures as drawn were too easy, that is, very few subjects made any errors save in the last segment of the figure and the range of scores was thereby so restricted as to greatly reduce the usefulness of the

test. While this last-named difficulty might have been overcome by redrawing the figures, the difficulty in overcoming the tendency to move the paper would remain. Therefore a substitute for this test called the Square Test was devised.

The Square Test consisted of a series of eight concentric squares with parallel sides. The direction of movement was from the outside toward the center, and the lines forming the outside square and the one next within and parallel made the channel of first degree in which the subject's pencil line must be drawn. An opening or gate led to the next inner passageway, which was narrower than the first. In a like manner each successive passageway or channel gave into another passageway which was narrower than itself. Thus the whole test consisted of seven passageways each more difficult to negotiate without error or contact than the one preceding it. (See Figure 1.) The test was *scored essentially the same way as the line test had* been. There was a power score or achievement score analogous to the score for distance achieved without error in the line test. Each square was divided into twelve equal segments. Of course the segments varied in length from square to square, becoming shorter as the squares became smaller and more difficult to negotiate without error. The power or achievement score for any individual would be the total number of segments successfully completed. The error score was simply the number of errors made, that is, the number of contacts made in completing the figure. This test proved much more satisfactory to give than the line test.

The last test of this original battery was the Circle Test. It was exactly the same in its fundamental plan as the Square Test. It consisted of a series of seven concentric circles with varying distances between the circles. The starting point of the test was at the center and the circles came closer to each other as one approached the periphery of the figure. Thus there were six channels or steps of increasing difficulty as one proceeded from the starting point at the center to the end of the test at the exit from the last or outermost circle. In scoring, each circle was divided into eleven segments. The segment which contained the gateway into the next circle was not counted so that the value of each circle perfectly done was ten and the total score was the sum of the segments completed without error. The error score was simply the number of errors or contacts made in completing the figure. There were two such series of circles in each test, in one of which the direction of movement was clockwise and the other counterclockwise. While some difficulty was experienced in having the children keep the test in one position all the time, the chief difficulty in administering the Circle Test came in making it clear which direction was to be taken in each of the two varieties of circles, clockwise and counterclockwise.

The total test battery of five tests took, in its original form, about 50 minutes to give. The children seemed to enjoy taking the tests very much indeed. Scarcely ever did we leave a classroom without having some child ask when we would come to give some more tests of the same kind.

The Handedness Score

The test scores so far mentioned have referred to the achievement of each hand separately. The problem of combining the scores of the two hands into an acceptable measure of the *handedness* remained to be met. Some studies had used the difference between the raw scores. However this did not seem to be a satisfactory procedure in this case since the calibration of the separate tests was so different that differences in raw scores would not have been comparable. A ratio of the left-hand divided by the right-hand score was another possibility, but this would have yielded a zero of unknown value whenever there was a zero in the numerator of the ratio. It would have other disadvantages also which may be more visual than real in a statistical sense. For example, a score of .25 would have meant that the score for the left hand was ¼ that of the right hand or that the right hand was four times as efficient as the left in that particular test (unless the score being dealt with was a score of errors, when the reverse would be true.) When the score for the *left* hand was four times as great as the score for the right hand, however, the ratio would be 4.00. This score of 4.00 would represent the same degree of left-handedness as a score of .25 would represent right-handedness, but the two look so different that it would be difficult to visualize this as being the case.

The scoring scheme finally decided upon may be expressed by the formula $\frac{R-L}{R+L}$ where R and L stand for the scores of the right and left hands respec-

tively. In the case of the error scores obtained in the last three tests of the battery the order of scores in the numerator was reversed $\frac{L - R}{R + L}$ so that a negative score always represented left-handedness. Using this formula the range of scores was limited to the values from minus one to plus one; a zero score means ambidexterity, that is, no choice between the performance of the right and the left hand. Any given score indicates the same amount of left-handedness as the same score would mean right-handedness if the sign were changed, that is, a score of $-.25$ means the same preponderance of the left hand over the right as a score of $+.25$ would mean a preponderance of right hand over left.

Expressed as a definition, the handedness score used in this study is the proportion that the difference in achievement of the two hands is to the total achievement on a given test, represented by the *sum* of the scores of the two hands.

The computation of these ratios for a large number of cases where a battery of four or five tests is used represented no small task. In the hope that some short-cut could be found the writer took this problem to Dr. Jack W. Dunlap, co-author of the *Handbook of Statistical Nomographs, Tables, and Formulas* (3). The result was the nomograph which is reproduced in simple form in the Appendix. A larger form, operated by two people, very quickly gave the desired values with any numbers up to 300. It proved to be a very great time saver and the writer is very glad to

acknowledge his indebtedness to Dr. Dunlap for this contribution.

The method of scoring the Line, Square, and Circle Tests to get a so-called "power" or achievement score was used because a small proportion of the children made "no-error" or perfect scores with the dominant hand. The resulting zero score when combined with the score for the other hand in the formula for the handedness ratio always gives a ratio score of 1.00, either plus (right-handed) or minus (left-handed). Strictly interpreted a ratio of one should mean almost complete inadequacy of one hand with very unusual superiority of the other. However, when a zero score with one hand was paired with a score of one or two errors for the other hand the resulting ratio of 1.00 entirely misrepresents the facts. What should be considered logically as practically equal achievement on the part of both hands is represented by a score which stands for exactly the converse, i.e., extreme preference of one hand over the other. A score of no errors should be a very rare occurrence. Fortunately, it is relatively rare. However, whenever it does occur, there is practically a misrepresentation of fact in the resulting ratio of 1.00 except when the error score with the other hand is fifty or more. This may be seen from the following illustration if it is granted that a score of zero errors and a score of one error do not represent very different amounts of skill. A score of one error with the right hand and a score of fifty errors with the left hand yield a ratio of .96.

$$\left(\frac{50-1}{1+50}=\frac{49}{51}=.96.\right)$$

A score of one error with one hand and a score of 100 errors with the other yield a ratio of .98; and it requires a score of one error with one hand and 200 errors with the other to yield a ratio of .99.[1]

No Reliability Coefficients for Preliminary Tests Available

The tests of this first battery were not planned to permit the computation of reliability coefficients by the split-halves method, since the original intention was to retest the same group with the same battery and thereby get an estimate of the reliability of the tests. The first testing was done in the late spring, shortly before the schools closed for the summer. In the fall the pressure of work in the school system was so great that the superintendent in charge did not feel justified in permitting so much additional time to be taken with this retesting where the immediate benefits to the school or to the children were felt to be so negligible. Since revisions of various kinds had been suggested by the experience gained with the tests, it was not felt that a testing and retesting with the old battery in a new locale for the sake of getting an estimate of the reliability of the battery was worth while. It was possible to get a reliability coefficient of the Target Test by the split-halves method and this, although lower than at first hoped for, was high enough to justify further experimentation. No attempt was made to

[1]In dealing with the handedness ratios throughout the remainder of this study decimal points have been omitted and the ratios have been treated as whole numbers ranging from —100 to +100.

validate the battery in this preliminary form with any other criterion such as a questionnaire study of hand preference or teachers' judgment of handedness. There were six children out of the 194 who completed all five forms of the test who were habitual left-handed writers. A composite score for the five tests of the battery, weighted upon the basis of the standard deviation of each test so that each test had a weight of one in the total, gave left-handed scores for each of the six. However, the number of cases is so small that no great importance is attached to this.

ANALYSIS OF OYSTER BAY DATA

The experimental population used in trying out these preliminary forms of the handedness test battery was comprised of eight classes in the Oyster Bay Public Elementary School. The following gives the distribution of classes and the number of children in each class:

Grade	4A	5A	5B	5C	6A	$6A_2$	6B	6C	All grades
Number of cases	23	23	32	27	26	25	25	23	204

The school is of very high average in respect to the progressiveness of the school plant and the methods of instruction used. Cooperation from school officials and teachers in administering the tests was all that could be asked for.

The data presented in the remainder of this section are included chiefly to afford a comparison between the showing of the tests in the preliminary form with that obtained with the later revised forms and is not intended to be used as the basis for conclusions better drawn from the more extensive data later collected.

TABLE 1

DISTRIBUTION OF HANDEDNESS RATIOS FOR PRELIMINARY BATTERY

		Pin	Target	Circle power	Circle error	Line power	Line error	Square power	Square error
	96-100				2				1
	91- 95				1		4		1
	86- 90				7		0	2	1
	81- 85				6		6	3	7
	76- 80				10		6	0	4
	71- 75			2	18		5	5	5
	66- 70			0	21		11	0	5
	61- 65			0	12		11	2	4
	56- 60			0	21		14	5	17
	51- 55	1		1	17		16	4	21
Plus	46- 50	0		2	16		20	8	18
	41- 45	4		5	19		19	5	21
	36- 40	6	1	9	8		24	7	12
	31- 35	14	0	8	6	3	14	7	26
	26- 30	27	2	20	4	7	8	10	9
	21- 25	39	3	24	4	13	12	12	11
	16- 20	38	8	13	3	22	10	14	6
	11- 15	25	18	24	4	51	6	23	3
	6- 10	14	66	26	2	46	2	26	4
	1- 5	15	60	21	0	35	0	18	1
	0- 4	5	12	7	1	9	1	11	2
	5- 9	0	2	8	0	2	1	13	2
	10- 14	0	0	2	0	5	0	1	1
	15- 19	2	1	1	1	1	1	4	1
	20- 24	0	1	1	0		1	2	2
	25- 29	3			1		0	2	1
	30- 34	1			1		1	0	1
	35- 39				1		0	0	3
Minus	40- 44				0		0	2	0
	45- 49				0		0	2	0
	50- 54				1		0	2	1
	55- 59				3		1	0	2
	60- 64							1	0
	65- 69							1	0
	70- 74							0	1
	75- 79							0	
	80- 84							1	

Table 1 gives the distribution of the Handedness Ratios for each of the five tests. For the Line, Square,

and Circle Tests the results of both the error and the power method of scoring are shown.

TABLE 2
HANDEDNESS TEST BATTERY GIVEN IN OYSTER BAY PUBLIC SCHOOL
N=194

Test	Median	Mean	Sigma
Dotting (Pin)	19.7	18.42	12.7
Target	7.1	7.40	8.0
Line power	10.4	10.39	8.6
Line error	44.1	44.31	22.5
Circle power	16.6	16.84	14.3
Circle error	56.6	52.18	27.0
Square power	12.9	16.32	27.9
Square error	42.4	37.12	28.3

It is obvious from Table 2 that there is considerable difference from test to test in the variability or spread of the scores. The S.D.'s given in Table 2 strengthen this impression, varying as they do from 8.0 in the Target Test to 28.3 in the Square Test, error method of scoring.

The means of the distribution also vary widely from the 7.40 of the Target Test to 52.18 for the Circle Test, error method of scoring. All the mean handedness ratios are positive but there is a wide difference in the amount of hand dominance represented by a mean score of 7.40 and mean score of 52.18. A certain amount of this difference is due, no doubt, to the nature of the scoring methods since two different methods of scoring are used.

Comparing medians and means we find that three of the distributions, those of the Dotting Test, the Circle Test, error scoring and Square Test, error scoring, are negatively skewed, i.e., the medians are larger than

the means and only in one test, the Square Test, power scoring, is there a noticeable skewness in the positive direction.

Since the reliability of these tests is unknown and the number of cases, 194, is relatively small it would not do to stress these comparisons very much. It will be interesting to note, however, to what extent the impressions given here are borne out in the later forms of the tests with larger number of cases.

TABLE 3

INTERCORRELATIONS BETWEEN HANDEDNESS TESTS IN PRELIMINARY BATTERY GIVEN IN OYSTER BAY PUBLIC SCHOOL

N = 194

	Dotting	Line power	Line error	Circle power	Circle error	Square power	Square error
Target	.290	.221	.100	.123	.203	.169	.142
Dotting		.237	.143	.114	.286	.149	.236
Line power			*.497*	.194	.251	.229	.336
Line error				.175	.311	.187	.363
Circle power					*.347*	.060	.127
Circle error						.212	.302
Square power							*.362*

Table 3 gives the intercorrelations of these original tests. The three coefficients which are italicized are the correlations between the error and power method of scoring the Line, Square, and Circle Tests. It is significant to notice that these three coefficients are higher than most of the other coefficients given and that the error method of scoring the Line, Square, and Circle Tests shows higher intercorrelations on the

whole than the intercorrelations between the Target and the Dotting Tests with the rest of the battery. Here again, the fact that the reliability of the tests is unknown greatly reduces the value of any conclusions that might be drawn from the intercorrelations given.[2]

[2]The casual reader may find it helpful, either at this point or at the end of Section III, to turn to the last section, which summarizes the specific aims of the study and the degree to which they have been achieved, and makes suggestions for further investigations along this line.

III.

REVISION OF THE TEST BATTERY AND ADMINISTRATION OF THE REVISED BATTERY

Nature of Test Revisions

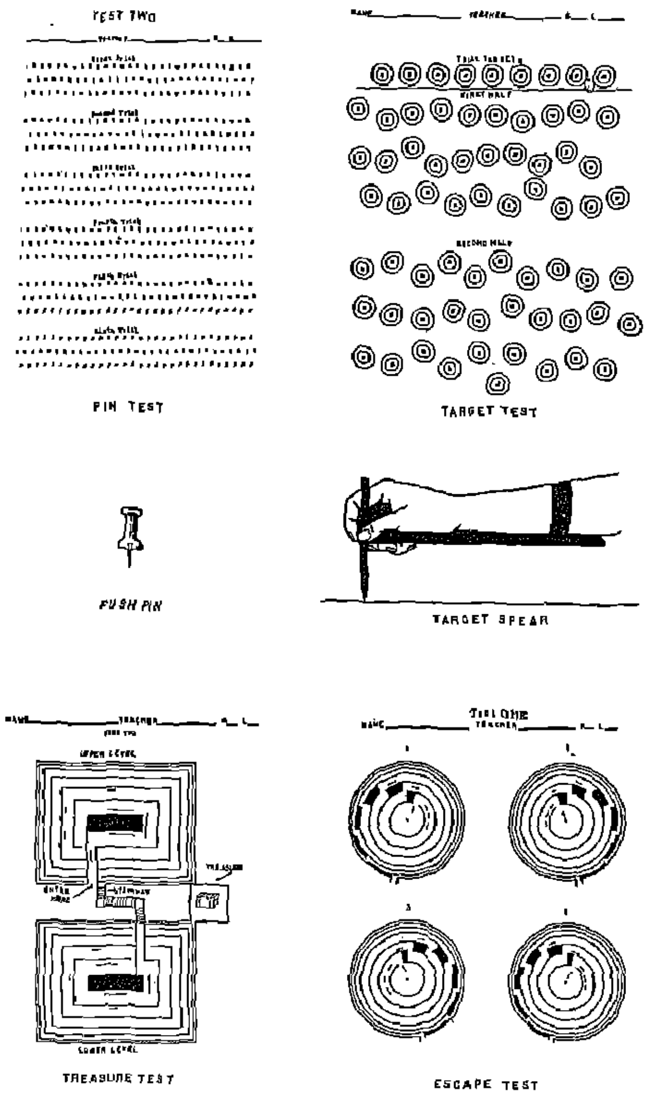
The administration of the tests in their original form at Oyster Bay revealed no administrative difficulties invalidating this method of testing with the exception of the Line Test, which was discarded. The revisions in the tests and in the administrative procedure indicated by this preliminary experience were largely in the nature of refinements of the tests and test directions and changing the forms of the tests slightly to permit the calculation of reliability coefficients from split-halves. The nature of the tests is such that this procedure seemed to be the most satisfactory method as well as the most convenient way of obtaining an estimate of reliability.[8]

Figure 2 shows the tests in their final form. The Treasure and the Escape Tests were reproduced by photographic offset process to give clearer definition of line, and the others mimeographed.

In the case of the Target Test the revision consisted chiefly of greatly increasing the length of the test so that it includes not twenty but sixty targets. There are ten targets to a line and the first half and second half each have three lines. In practice the two halves were separated by a short rest period which was als

[8]See (1) for excellent discussion of the split-halves method of obtaining reliability coefficients.

FIGURE 2



HANDEDNESS TESTS — REVISED FORMS

valuable because it permitted redirection in the cases where misunderstanding persisted as to the procedure to be followed. The test directions were not altered except to emphasize more strongly the points at which misunderstanding had previously been noted. A metronome was introduced so that the counting for this test would be uniform for all classes.

For the Pin Test the revision was also one of lengthening the test greatly. From 150 circles arranged as one continuous unit, the number of circles on the page was jumped to 600 in six trials of 100 circles each. The time allowed for each trial was 45 seconds, determined by giving the tests to a small group of adults so as to ascertain the longest possible time which could be allowed for each trial with a minimum of completed tests. This time limit proved to be very satisfactory in practice. Only a comparatively small number of children ever completed a trial with anywhere near perfect accuracy. In this test there was a question whether the factor of speed was to be stressed at the expense of accuracy or vice versa. Nothing in the preliminary testing threw any great light upon this point so the two factors were emphasized just as equally as possible and the final decision as to the procedure in scoring was left to be made on the basis of the relative reliability coefficients obtained by the two methods, i.e., by counting only the responses where the pin puncture was definitely on the line or within the circle versus counting the total number of responses made.

In the place of the corrugated cardboard used to mount the Pin and Target Tests in the preliminary

phasis in the test directions upon the fact that there were to be three shots at each target took care of most of the cases where misunderstanding might have occurred.

In the Pin Test, because of the simplicity of the procedure, it was not felt necessary to have any trial period and this conclusion was justified by the experience of giving the tests. This was by far the easiest test to administer, but the one which was liked the least by the children.

A trial test was not possible in the case of the Treasure and Escape, due to lack of space on the paper and also the lack of time. To offset this, large copies of these tests were made by a photostatic process (30 x 20 inches) which were used to demonstrate the exact procedure. This procedure was adopted after trying a large hand-drawn duplicate upon which the procedure was illustrated without any lines being drawn. However, this proved to be much less effective than the large copies upon which the children could actually see the examiner doing the thing which he required of them.

Problem of Finding Experimental Populations

The problem of the locale in which the experimentation was to be carried on for the larger number of cases necessary after the preliminary forms of the test had been given, analyzed, and revised, was, as it usually is, a vexatious one. The schools within the New York metropolitan area have certain very definite advantages and also certain very definite disadvantages so far as research of this type is concerned. They have the advantage of nearness and of being, in many cases,

research-minded enough to welcome experimentation even though they do not reap an immediate benefit from the work done in the schools. They have the grave disadvantage more often than not of being highly selected in one or more directions. For example, a school in or near Harlem or certain parts of Brooklyn may have a very large proportion of its population drawn from the colored race; or a school in a selected residential district will have a disproportionately large share of highly gifted children. On the other hand, schools outside of Metropolitan New York, which have a more heterogeneous and balanced population, are often trying out various new methods of their own involving research, or are more deeply involved in other outside affairs so that they do not welcome outside experimentation. The three schools in which this work was finally carried out were within the metropolitan area and shared the disadvantage in differing degrees of having a large proportion of Negro and foreign pupils. In one school, the average intelligence was distinctly below normal, and in the other two approximately normal. Another disadvantage in testing in the New York City schools is the crowded conditions often encountered, but the splendid cooperation of both administration and pupils, in practically every case, did much to minimize this difficulty. In P.S. 42, Brooklyn, thirteen classes in grades 4, 5, and 6 were tested; in P.S. 44, Brooklyn, ten classes in grades 4 to 8, inclusive, were tested; and in P.S. 43, Manhattan, twelve classes in grades 7, 8, and 9 were tested.

SCORING THE TESTS

In scoring the tests a good deal of trial and error experimentation had to be resorted to in order to determine the most effective procedure.

In the Target Test the plan of assigning a score of three to the bull's-eye, two to the next ring, and one to the outer ring was used, both in the preliminary test and in the final test. The chief problem here, which seemed beyond administrative control, was to regulate the strength of the blow with which the child hit the target board. There seemed always to be a small minority who would strike with such force that the wooden part of the stylus would be driven into the target board, while, at the other extreme, there was a small minority who would strike so lightly that the abrasion made by the target spear would be hardly perceptible. To aid in scoring the tests where the abrasions were faint, a mimeoscope, such as is used in making tracings or cutting stencils, was used effectively. To aid in scoring the Treasure and Escape Tests, a large reading glass was mounted on a flexarm, such as one finds in a goose-neck desk lamp, which was, in turn, mounted upon a wide board. This permitted the adjustment of the glass at any angle and greatly facilitated the inspection of these tests while diminishing the eye strain necessarily attendant upon such minute work as checking and counting the number of times the line drawn by the child touched the printed line of the test.

In the case of the Pin and Target Tests there was the problem of determining the two halves of the tests

in such a way that they would be as nearly as possible equivalent forms. For the Escape and Treasure there was the problem of the method of scoring to be used. In II the so-called power method of scoring was described. This was essentially a method of counting units or segments of work accomplished without error. In the case of the Treasure Test some arbitrary unit, such as the centimeter or some fractional part of an inch, might have been used without much difficulty, but this would have been very difficult in the case of the Escape Test. Chiefly, therefore, to have a type of measuring unit common to both tests, the proportional distance method was devised. By this method, each level of difficulty for both the Treasure and Escape Tests was divided into ten equal units. These units varied in absolute length while being alike in proportional length with any particular level of difficulty as the reference point. For the easier levels of difficulty, the absolute length of the unit was less, tending slightly to equalize the chances of making an error from the easier levels of difficulty to the harder. This tendency appeared to be good since a cursory examination of the completed tests showed that by far a majority of the errors made came in the last or highest level of difficulty. Having decided upon the basic method of scoring, the next problem was to choose the point of reference to which the scoring plan would be directed. If the number of perfect segments up to the *first* error were to be taken as the score, the chance errors would be likely to be large. Therefore, as one point of reference, the second error made by the child was taken and

the number of segments up to but not including that segment which contained the error was called that child's score on that test. Lest this should not give great enough stability, the median error was also tried out as the point of reference. The obvious method of counting the number of errors made was the third tried out. The chief drawback of this last-named method was that a score of no-errors in one hand gave a misleading ratio of handedness in many cases. (See Section II.)

Scoring Methods Tried Out

Having these various scoring possibilities to test out, three classes from the first school tested were selected as a group for preliminary study. Classes with mean I.Q.'s as near as possible to 100 were chosen in order

TABLE 4
Reliability Coefficients for Limited Number of Cases in Grades 5 and 6 According to a Variety of Scoring Methods

Test	Uncorrected	Corrected	Index of reliability	N
Target Test				
Lines 1, 2, 3 *vs.* 4, 5, 6	.519	.683	.826	119
Lines 1, 3, 5 *vs.* 2, 4, 6	.594	.745	.863	119
Pin—all	.682	.811	.901	112
Pin—accuracy	.717	.835	.914	112
Escape Test				
Power—median error	.130	.230	.480	109
Power—2nd error	.307	.470	.684	109
Number of errors	.511	.677	.807	109
Treasure Test				
Power—median error	.060	....	....	112
Power—2nd error	.271	.426	.653	112
Number of errors	.438	.609	.780	112
Treasure—Escape Test				
Power—2nd error	.437	.608	.780	101
Number of errors	.513	.679	.823	101
Composite scores	.701	.824	.908	97

that intelligence might not be a disturbing factor. This gave around a hundred cases, the number varying slightly from test to test. Table 4 gives the reliability coefficients for each test scored by the variety of methods outlined above. From this table it will be seen that, in the case of the Target Test, lines 1, 3, and 5 *vs.* 2, 4, and 6 make a better division of the test in the sense that they yield a higher reliability coefficient. Therefore this method of division was chosen.

In the case of the Pin Test the method of counting only those responses which were accurate in the sense that the puncture was either on the line or inside the circle gave the higher reliability coefficient, although the difference was small. This method of scoring the Pin Test was therefore adopted. In either of these cases, the differences were not so great as to cause the elimination of the method if there had been other good reasons why it should have been adopted. Such reasons were not found, however, so the size of the reliability *r*'s was taken as the criterion.

In the case of the Treasure and Escape Tests, it is obvious that there is a great difference between the reliability of the various methods of scoring tried out. The power-score method proved so definitely unreliable that if this had been the only method of scoring these tests they would have had to be thrown out. The number-of-errors score was distinctly better than either of the power scores and compared favorably with reliabilities obtained with the other tests. All the reliability coefficients were lower than hoped for, but were felt to be high enough when corrected for the

full length of the test by the Spearman-Brown Prophecy Formula to justify the use of the tests as group measuring instruments to yield mass data, but not valuable for individual diagnosis. It was hoped that higher reliability coefficients would be obtained when more, and more typical or more normal populations were tested. The very low reliabilities obtained with the power method of scoring probably were the fault of the tests rather than the scoring scheme since they came from the greatly restricted range of scores obtained by this method of scoring. If the levels of difficulty below the last level had been harder so that errors would have been more widely distributed instead of being concentrated in the last level of difficulty, this method of scoring would have been much more reliable, and probably quite satisfactory.

Since the Treasure and Escape Tests seemed on the surface to be very similar in the thing or skill they measured, it was decided to try combining the two tests into one test to see if this would yield a higher reliability coefficient. Table 4 shows that this coefficient was only slightly larger than that obtained for the Escape Test alone, so it was concluded that a combination of the two tests would be of doubtful advantage.

It was necessary, however, to combine all scores for all the tests of the battery into one composite score which could be taken as the final handedness score for each individual, to be used especially when it was desired to epitomize the handedness of an individual in relating it to other variables, such as intelligence, school

marks, etc. Since each test was as valid a measure of handedness as each other test, so far as could be seen on the surface, the best composite score would be one which was influenced equally by all the tests. Therefore, since the standard deviations varied widely, the tests were weighted roughly in proportion to the size of their S.D.'s. By this criterion the Treasure and Escape Tests each had a weight of one, the Pin Test a weight of two, and the Target Test a weight of four. The reliability coefficient of this composite score for the 97 cases, where the tests were complete, was found to be .701, uncorrected, a value distinctly higher than the reliabilities of the several parts from which it was derived, with the exception of that of the Pin Test.

Reliability Coefficients for Larger Groups

Having decided upon the methods of scoring to be used, the remainder of the tests was corrected and reliabilities were calculated for a total group of 18 classes, involving six grades from four to nine, inclusive, three classes to each grade.[4] These data are presented in Table 5. From this table we see that all the corrected reliabilities are higher than .80 with the exception of the Treasure Test. The Index of Reliability, an estimate of the highest reliability which can be expected from any test in its present form, is higher than .90 for every test but the Treasure Test, for which it has a value of .87. These results were

[4]In a later article the reliabilities will be presented separately for each grade of the six grades studied here.

TABLE 5
RELIABILITY COEFFICIENTS OF EACH TEST OF THE BATTERY AND THE COMPOSITE SCORE**

Test	Uncorrected	Corrected*	Index of reliability	N
Target Test	.7024	.8252	.9084	668
Pin Test	.8102	.8951	.9461	651
Escape Test	.7296	.8437	.9185	619
Treasure Test	.6115	.7589	.8711	632
Composite score	.8330	.9089	.9534	651

**Reliabilities for a grade range from four to nine inclusive, three classes for each grade.

*Spearman-Brown Prophecy formula used to predict reliability of whole test from the correlation of halves.

considerably better than expected on the basis of the trial reliabilities for 100 odd cases discussed above and were considered satisfactory for the purposes for which the tests were devised.

FURTHER DISCUSSION OF TEST VALIDITY

The question of the validity of the tests was discussed at some length in Section II when the preliminary forms of the tests were described. The extent to which a test offered equal opportunity to either hand without bias was suggested as the most satisfactory criterion of validity. The degree of correlation with some other measure of handedness was considered as of very doubtful value since no satisfactory measures of handedness were known. The criterion of equal opportunity should apply to:

1. Equipment.
2. Process or procedure involved in taking the test.
3. Physical environment in which the tests were taken.
4. Mental hazard.

Of course it is impossible to reduce the degree to which the tests meet the criterion of equal opportunity for each hand to anything so objective as a coefficient or ratio, but an analytic discussion will be undertaken with the hope that an unbiased judgment may be reached.

1. *Equipment.* The equipment used in the Target and the Pin Tests was entirely novel to the children and was as well adapted to one hand as to the other. In the Treasure and Escape Tests the use of a pencil raises the question whether the effect of the writing habit might not be especially strong in favor of the writing hand. There can be no doubt that the writing habit did affect these two tests somewhat; however no claim is made in this study that handedness was measured *apart* from the influence of environment, but simply that existing handedness, due either to innate disposition or to habit patterns, was reduced to a ratio which would be relatively the same for equal amounts of left-handedness and right-handedness, although opposite in sign. This would permit the plotting of a continuous curve with all degrees of handedness represented. It is felt that it is just as impossible to separate the effect of environment from the effect of inheritance in the measurement of handedness as it is in the measurement of intelligence, and any information to be obtained concerning the inheritance of handedness or laterality must come from the study of growth curves and the study of laterality in animals.

2. *Process.* In the Target and Pin Tests the direction of movement was from right to left for both the

right and the left hands. There has been some discussion in the literature to the effect that the natural movements of the hands are excentric, but not much evidence has been produced to substantiate this claim. On the contrary, through long-established habits, both the left- and the right-handed are accustomed to work on paper from left to right, and it was felt that this direction of movement would interfere least with the achievement of either the left or the right hand.

3. *Physical Environment.* The fact that the classrooms in which the tests were given were nearly all of the conventional type with fixed seats and left-wall lighting made it impossible to make any adjustments as to lighting for the left-handed children. Therefore, it is possible that the left-handed were under some slight handicap in this respect, although the importance of the handicap does not appear to be very great. Since their dominant hand was the one handicapped, the effect would be to decrease the score of this hand while the right and non-dominant hand, working under optimum conditions, might do better than expected. This would tend to bring the scores for the two hands nearer together and reduce the handedness ratio in the direction of ambidexterity. For the right-handed, however, the poor lighting would handicap the non-dominant hand, increase the difference between the scores, and make for a larger handedness ratio. It was not possible to evaluate the importance of this factor experimentally, but it was felt that the desire to accommodate or overcome this lighting handicap would probably offset the effect of bad lighting sufficiently to permit us to ignore it.

4. *Mental Hazard.* For some the attempt to do anything calling for relatively close coordination with the non-dominant hand resulted in such a low degree of success compared to the success achieved in the same task with the dominant hand, that emotional blockings occurred. This seemed especially true with children of low intelligence. Usually, however, a word to the effect that this was a test intended to show up the differences between the hands was sufficient to relieve this difficulty. Although all four tests were somewhat affected by this factor, the Escape and Treasure Tests were the ones which most often caused it.

Effect of Vision Unknown

The effect of vision upon the test results is an unknown factor. Visual acuity records were taken from the school data but are probably so unreliable that they are of little value. However, these data are to be analyzed later and the results reported in a separate article. A cursory examination indicates that the high proportion of 20/20 vision in both eyes will make this a negligible factor so far as its effect upon the statistics reported in this monograph is concerned.

IV.

THE LEFT-RIGHT CRITERION QUESTIONNAIRE

THE PREFERENTIAL METHOD OF MEASURING HANDEDNESS ANALYZED

Many, if not a majority, of the handedness studies made in this country have attempted to measure handedness by determining the preferred hand in a series of performances such as throwing, reaching, writing, etc. Strength of preference has been determined by the proportion of those acts done with the dominant hand.

In some cases the actual performances have been used in testing for hand preference but in many cases some kind of questionnaire has been the means of securing this information. Such a method has the advantage of being a very easy way of collecting data but suffers from some very real disadvantages. While strength of hand preference may be roughly determined by the proportion of acts performed with the dominant hand such a proportion could never be considered an adequate or precise measurement for several reasons. In the first place, the major assumption, that any performance involving a choice of hands is equally as valid as a measure of hand preference under all conditions as every other performance involving such a choice, is obviously false. Carrying a weight, for example, can not be as good a measure of hand preference as throwing a ball at a target because success in the first case requires only strength enough to carry the object

and the difference in strength between the hands is so slight that if a person were able to carry an object in one hand he could, with very little less ease, carry it with the other if there was any reason to do so. On the other hand, throwing a ball with accuracy at a target involves such elements of skill and precision that the person who could do as well with one hand as with the other would be very exceptional. In the second place, the material set-up often predisposes one toward the use of the right hand even though one is naturally inclined to use the left. For example, even left-handed people shift gears with the right hand because the gear-shift lever is placed at such a disadvantage that the left hand could be used only with the greatest difficulty. It may be said that this is an extreme example and that such a question would not be used in a scale intended to measure hand preference. Another example may serve to illustrate the fact that this differs only in being more extreme than many questions which are used. For instance, it is not at all uncommon to find the question "With which hand do you throw a ball?" included in such a scale, and yet it is very difficult for a left-handed ball player to buy a glove to fit his right hand so that his left will be free for throwing. In the third place, the immediate situation may influence the choice of hands, when under different circumstances a different choice would be made. For example, if one were shaking hands with a man who had lost his right arm one would be likely to use the left hand although in almost any other situation the right hand would be used.

The illustrations used may not be especially fortunate but will serve if the point is made clear that, as a method of getting a precise and accurate estimate of the strength of manual dominance or superiority, the method of using the proportion of a series of acts performed with the dominant hand is very unsatisfactory.

Development of Handedness Criterion Valuable By-Product

On the other hand almost any other method that has been suggested in the literature to secure a quantitative handedness score involves such labor as to make its use well nigh prohibitive unless some means is hit upon to segregate those who have left-handed tendencies. In other words, to measure the handedness of a large unselected population when it is known before hand that better than 90% of them will be right-handed when the chief purpose is to study the ways in which left-handed or ambidextrous people are atypical is an expensive and wasteful procedure. Only in the one case where it is desired to study the distribution curves of handedness with an unselected population is such a procedure justified. A valuable by-product of such an investigation of distribution curves, however, should be the development of some easy method of selecting those individuals from a large random population who have left-handed or ambidextrous tendencies.[5]

[5]Since the effectiveness of the criterion which we propose to set up is to be tested against the findings of the test battery already described, it is necessary to consider now at what point the line shall be drawn dichotomizing our distribution curves into left-handed and

Such a criterion should have the following characteristics:

1. It should be as objective as possible.
2. It should require very little time to give and less to score.
3. It should be of known reliability.
4. It should be of known efficiency in selecting out those individuals who will fall into the ambidextrous and left-hand sections of the distribution of handedness when some more adequate methods of testing are applied.

Nature of Criterion Selected

It was for the purpose of developing such a criterion that the L-E-R Criterion Questionnaire was designed and the degree of success in thus accomplishing this purpose may be judged by the data presented in this section.

---

right-handed groups. This is necessary because our tests yield a continuous curve with all degrees of dominance represented. The obvious point at which to cut the curve is at the zero or no-handedness point and if the tests were 100% reliable this would be possible. Since they are not even approximately 100% reliable, it is necessary to allocate a range plus and minus from the point of zero handedness as a sort of no-handedness area or area of ambidexterity, and all the scores above this range will be considered right-handed scores, while all scores below this area will be considered as left-handed scores. In order that the choice of this range or area should not be purely arbitrary the probable errors of measurement were calculated, and the area roughly one P.E. of measurement plus and minus from zero is called the range of ambidexterity. This range is variable from test to test and is only approximate since the roughness of grouping makes it impossible to apply the criterion precisely. The data concerned are presented in Section V in connection with the discussion of distribution curves and will be discussed at greater length at that point.

The following ten questions were chosen, eight of which concern habitual unimanual acts ordinarily performed under such circumstances as would be least likely to influence the choice of hand used. The other two questions involve the subject's own judgment of his relative achievement in strength and in reaching, two measures of manual laterality very commonly used.

L-E-R Criterion Questionnaire

We want to find out which hand you usually or habitually use to do certain everyday things which we all do. In answering each question, put a circle around one of the letters to tell us which hand you most always use. R stands for Right Hand. E stands for Either Hand. L stands for Left Hand. For example, suppose the question is:

With which hand do you write? L E R

If I were answering that question I would put a circle around R—like this:

With which hand do you write? L E R

If I could write just as well with one hand as with the the other, I would put a circle around E which stands for *Either Hand*—like this:

With which hand do you write? L E R

If I wrote with my left hand, I would put a circle around L like this:

With which hand do you write? L E R

Answer each of the questions by putting a circle around *one* and only *one* of the letters. If you are not sure, make a guess.

1. With which hand do you hold a jack-knife or paring knife?	L	E	R
2. With which hand do you write?	L	E	R
3. With which hand do you throw a ball?	L	E	R
4. With which hand do you erase the blackboard?	L	E	R
5. With which hand do you shoot marbles?	L	E	R

6. With which hand do you usually brush your teeth?	L E R
7. With which hand do you draw or paint?	L E R
8. With which hand can you reach higher?	L E R
9. With which hand do you usually hold a glass or cup, when drinking?	L E R
10. *Which hand do you think is stronger?*	L E R

Considering the questionnaire in the light of the four desirable characteristics outlined above:

1. *Is it objective?* No questionnaire can ever be considered entirely objective since each question must receive a subjective answer. It can be objective, however, in the sense that the questions asked may be so definite and so unambiguous and so much a matter of fact, rather than opinion, that the answers to the questions will correlate highly with the performance in whose place they stand. With a much longer and more complicated list of questions Koch, *et al* (10) found a very high degree of agreement between actual performance and questionnaire response. To be sure their experimental group was made up of mature individuals and not children, but on the other hand their list of questions was much more complex and difficult to answer.

2. *Is it quickly given and scored?* The questionnaire requires less than five minutes of class time to give so does not interfere to any appreciable extent with the classroom program. It is also very simple to score. As seen in the reproduction of the questionnaire given above each question could be answered either left, right, or either, the last-named category being de-

signed to take care of those cases where the subject was not sure or felt that one hand might be used quite as much, or as well as, the other in the performance of a particular act. In scoring it was desired to use some method whereby the answer E would reduce the score but not as much as an R or L answer. Also it was felt that it would be advantageous if the score so obtained would be roughly comparable to the ratio scores used for the tests of the battery. The scoring formula used for the battery tests would not work in this case but a modification of it was finally adopted,

$$\frac{R-L}{\text{No. of questions}}$$

In this instance this reduced to R—L since decimal points were omitted in the actual computations and the number of questions was ten. Less than a minute was required to score a paper and after a little practice only a fraction of this time.

3. *Is the reliability of the questionnaire known?* The problem of determining the reliability of the questionnaire was approached in two ways. It was felt that the number of questions in the test was so small that the split-halves method would not do justice to the real reliability of the test. However, since the questions seemed to fall naturally into two rather comparable series of five each this method was applied. Questions 1, 2, 3, 9, and 10 composed one half and questions 4 to 8, inclusive, the second half. There were available 831 cases covering a range of six grades from 4 to 9 inclusive. The obtained reliability coefficient

corrected by the Spearman-Brown Prophecy Formula is .9185, a figure surprisingly high for such a short questionnaire. For fear there might be something spurious about this value the questionnaire was repeated with a group of 140 fourth-grade children after a lapse of ten days from the time they had taken the test the first time. The coefficient obtained with this group was .9701. Therefore it is felt that from the point of view of the conventional methods of measuring reliability the questionnaire is quite satisfactory for the purposes for which it was made.

4. *It should be of known efficiency in selecting those individuals who have left-handed and ambidextrous tendencies.* Herein lies the very great advantage of having tested a large group of subjects, previously unselected, with the handedness test battery, and also with the Criterion Questionnaire. By applying the Criterion to these data some reliable evidence can be obtained concerning what might be expected from the use of the Criterion for the selection of additional cases from new populations. The first evidence concerning the efficiency of the Criterion is the correlation coefficients between it and the tests of the battery and the composite handedness score. These coefficients were obtained with two groups of cases, one containing approximately 400 cases and the other approximately 300. The grade range was the same as above, i.e., 4 to 6 inclusive, for each group. The two coefficients for the Criterion with each test of the battery and for the composite score have been averaged by changing them into z-values according to Fisher's

tables (4, pp. 178-184), averaging, and reconverting these values into values of r.

These coefficients are as follows:

Criterion with Target Test	.3459
Criterion with Pin Test	.4985
Criterion with Treasure Test	.5214
Criterion with Escape Test	.5440
Criterion with Composite Score	.6339

It is obvious from these values that there is a very considerable degree of relationship between the Questionnaire and the tests of the Battery, although the relationship varies from test to test. In general it may be said that the relationship increases as the test requires finer and finer coordinations, and it is especially significant that the highest correlation is that with the Composite Handedness Score, since this is the score which best sums up or epitomizes the handedness of an individual so far as it is measured by these tests.

The correlation coefficients between the Criterion and the tests of the battery are not sufficient, however, to answer the question concerning the efficiency of the instrument as a selective agent for the purpose of choosing other, additional cases from a random population. The question may be stated this way: if the L-E-R Criterion Questionnaire had been used to select from the population included in this investigation those who had left-handed tendencies with what efficiency would this have been accomplished? Furthermore, what proportion of the cases so selected would have been found to be left-handed when tested with the battery? Before these questions can be answered some critical L-E-R score must be decided upon below which all

cases are to be considered as showing sufficient signs of being left-handed to be included in the retesting with the Handedness Battery. The selection of this score is more or less arbitrary. It should be a score such as would pick out a very large percentage of the left-handed or left-handed and ambidextrous, depending upon the scope of the proposed study. Yet it should not select for extensive additional testing too great a percentage of cases which would later prove to be essentially right-handed. Table 6 gives the results when two such critical scores are used, (1) zero, and (2) +6.

Considering first zero as the critical score we find that in the neighborhood of 80% of those who have left-handed scores (i.e., who are roughly below one P.E. of measurement of the zero score for each test in turn) would have been selected by the criterion for further testing from the population used in this investigation. If we desired to include the ambidextrous as well, the criterion would have selected between 55 and 65% of this wider range of cases with the exception of the Target Test where, due to the very high proportion of the subjects who got ambidextrous scores, the percentage drops to .1639. The proportion of these cases that would have proved to be left-handed on the test battery is in the neighborhood of 50% except for the Target Test, where it is lower, and if the wider range is used so as to include the ambidextrous the proportions profitably selected would be higher, ranging around sixty.

If a more inclusive criterion score is desired, i.e.,

Percentage of the left-handed* who have scores between —10 and +5 inclusive	.9047±.07 (21)	.9210±.04 (38)	.9428±.04 (35)	.8947±.05 (38)	.9473±.04 (38)
Percentage of the left-handed and ambidextrous who have scores between —10 and +5	.2704±.02 (366)	.7808±.05 (73)	.6718±.06 (64)	.7962±.05 (54)	.9183±.04 (49)
Percentage of those having L-E-R scores between —10 and +5 who are left-handed	.1216±.03 (148)	.2380±.03 (147)	.2426±.04 (136)	.2312±.03 (147)	.2535±.04 (142)
Percentage of those having L-E-R scores between —10 and +5 who are left-handed or ambidextrous	.4662±.04 (148)	.3877±.04 (147)	.3161±.04 (136)	.2925±.04 (147)	.3169±.04 (142)
Percentage of the left-handed who have scores between —10 and —1	.8095±.09 (21)	.8421±.06 (38)	.8000±.07 (35)	.7895±.07 (38)	.8684±.05 (38)
Percentage of the left-handed and ambidextrous who have scores between —10 and —1	.1639±.02 (366)	.6575±.06 (73)	.5469±.06 (64)	.6396±.06 (54)	.8163±.06 (49)
Percentage of those having L-E-R scores between —10 and —1 inclusive who are left-handed	.2698±.06 (63)	.5079±.06 (63)	.5088±.07 (57)	.4688±.06 (61)	.5484±.06 (62)
Percentage of those having L-E-R scores between —10 and —1 who are left-handed or ambidextrous	.9524±.03 (63)	.7619±.05 (63)	.6140±.06 (57)	.5574±.06 (61)	.6452±.06 (62)

*I.e., have scores below the lower limit of the standard error of measurement applied to zero.

if it is necessary to select a higher proportion of those who have left-handed or ambidextrous tendencies at the cost of including a greater proportion of cases who will eventually prove to be right-handed, according to the test battery, the score +6 would seem to be logical value to choose. That is, by this criterion, all those having scores between —10 and +5 inclusive would be retested. A score of +5, the highest value in this range, could be obtained in three ways according to the scoring formula, $\frac{R-L}{10}$. One combination yielding this score is 7 R's, 2 L's and 1 E. Another would be 6 R's, 1 L, and 3 E's and the last is 5 R's and 5 E's. Any one of these combinations suggests that the right hand is the dominant hand but with unmistakable tendencies in the opposite direction. The result of the application of this more rigid criterion to the data of this investigation is also given in Table 6. We see that in practically every case better than 90% of the left-handed have L-E-R scores within this extended range. If we include both the left-handed and the ambidextrous the percentages drop with an average between 75 and 80%, if the Target Test is not considered. In other words, if the —10 to +5 criterion is used, we might expect to select about 90% of those in the whole population tested who would have been left-handed if the test battery had been given to the whole group, and if we wanted to include the ambidextrous we could expect to select about 80% in this way. However, of those selected by the application of this criterion only

23 or 24% would be left-handed when tested with the battery. When it is considered that the test battery takes a considerable time to give, and longer to score, to have three-quarters of the group prove to be right-handed is a rather high price to pay for the additional cases secured by the use of the more rigid criterion.

Which criterion should be used in selecting additional cases from another population is dependent upon the purposes for which the cases are needed. A cursory examination of the correlation charts suggests that any other criterion score between 0 and +6 would not help enough to be worth while and certainly the number of cases which would be tested to no profit would prohibit taking a higher value than +6.

To summarize this section it may be said that a criterion has been developed for the selection of additional cases of left-handedness from an unselected population that will pick out approximately 90% of those who are left-handed if the more rigid selection is made or approximately 80% if the less selective method is used. This Criterion should be of great value in shortening the work of any additional investigations concerning the nature of any differences which may exist between left- and right-handed children.

V.

DISTRIBUTION CURVES OF HANDEDNESS

Distribution Curves Important

The problem of the distribution of handedness is one of the most important and one of the most difficult and illusive to treat of any of the various related aspects of this study. It is intimately bound up with the definition of handedness with which the investigator starts and the method of measurement used. It has peculiar difficulties of its own beyond those ordinarily found in curve fitting. The necessity for large samples, always present if reliable results are to be obtained by ordinary methods of curve fitting, becomes increasingly important in considering the distributions of handedness, since it is only when a very large sample is used that the influence of the left-handed section of the population upon the total can be reliably assayed.

Why Important?

Some of the difficulties met suggest the principal reasons why the problem is important. It has been rather naively assumed by many investigators in this field that any measure which can be used to discriminate left-handedness and right-handedness is as good, that is, as reliable and valid, as any other measure. Further, there has been the underlying assumption that all measures of handedness are positively correlated to the point of reliable prediction, that is, reliable generalization from a single measure to a "true" measure of handedness. Thus we have the rather ridiculous situa-

tion of investigators measuring the handedness of a small sample by some rather crude method of unknown reliability and from such measure attempting to generalize about the relation of handedness to other traits of character and personality as well as other characteristics of mental organization. Therefore the most fundamental problem in the study of manual laterality *after* the tests have been shown to be *reliable* and *valid* is the study of the way handedness distributes itself on such a variety of tests, measuring different aspects of the same trait. From such a series of obtained distributions the following important questions may receive consideration:

1. Will each test in a series of handedness tests, practically equal in objectivity, reliability, and validity, yield the same general type of curve? For example, tests of intelligence of equal validity and reliability may confidently be expected to yield a normal curve, although the items comprising the tests differ greatly.

2. If one particular curve type is found to fit all distributions of handedness, does that curve type have any peculiar properties which will facilitate its use statistically?

3. If a variety of curve types is found, what characteristics are common to them all, i.e., are they all skewed; is the skewness uniform in direction; is the degree of kurtosis present the same for all curves?

4. In what way do the types of curves found reflect the particular properties of the tests used?

Hypotheses Concerning Distribution Curve of Handedness

Several hypotheses concerning the type of curve which may be expected to describe handedness have been suggested or implied in the literature. At one extreme are those who think that handedness may be described by a two-category curve, i.e., all persons either right-handed or left-handed with no intermediate degrees of handedness and true ambidexterity nonexistent. A modification of this extreme would be a continuous U-shaped curve with the antimode representing ambidexterity, but not necessarily symmetrical so far as the height of the ordinates representing the extremes of leftness and rightness are concerned. Or handedness, many feel, can be described by some I-shaped curve of unknown skewness and kurtosis. A special case within this rather large grouping would be the normal curve so often found or assumed in psychometric investigations.

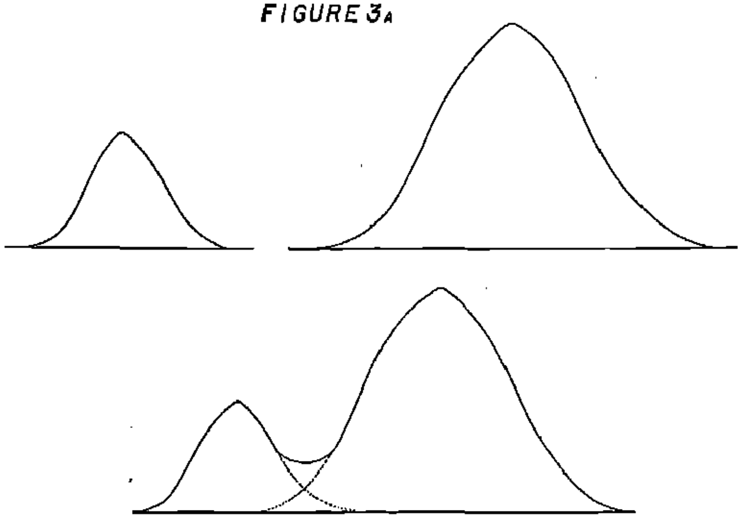
A third possibility that has aroused a great deal of interest on the part of some investigators is that handedness might logically be expected to be bimodally distributed with one mode representing the central tendency of the left-handed group and the other, much greater mode, representing the central tendency of the right-handed group. Some slight evidence from one or two American investigations bears on this point but is too unreliable to justify any general conclusion to this effect. A true bimodal distribution is rarely found. In most cases what appears to be bimodality eventually turns out to be due to the combination of

data from two different but homogeneous populations. Without denying the possibility of obtaining a distribution of handedness for some particular test that would have a reliable second mode for the left-handed population, it must be shown that the handedness population is homogeneous before it can be said that handedness is truly bimodally distributed. Homogeneity would seem to be a matter of degree. Very rarely would one obtain a distribution that would be entirely homogeneous. For example, a distribution of intelligence quotients in certain schools of New York City would include those of a large proportion of Negroes. (This is true of two of the schools used in this study.) If there were a genuine difference in the intelligence of Negroes and whites, such a distribution would not be homogeneous. However, the difference, if present, is so slight and the resemblances between the two distributions so much more important than the differences that the combined populations yield a distribution approximately normal, and not greatly atypical in any respect. If a distribution of the intelligence of man and a distribution of the intelligence of the great apes could be combined, in all probability this distribution would be sufficiently atypical to suggest immediately that two heterogeneous groups had been put together.

This discussion suggests a possible hypothesis to be considered along with the rest, namely, that any distribution of handedness really contains two distributions, one of the left-handed and one of the right-handed. According to this theory any distribution of

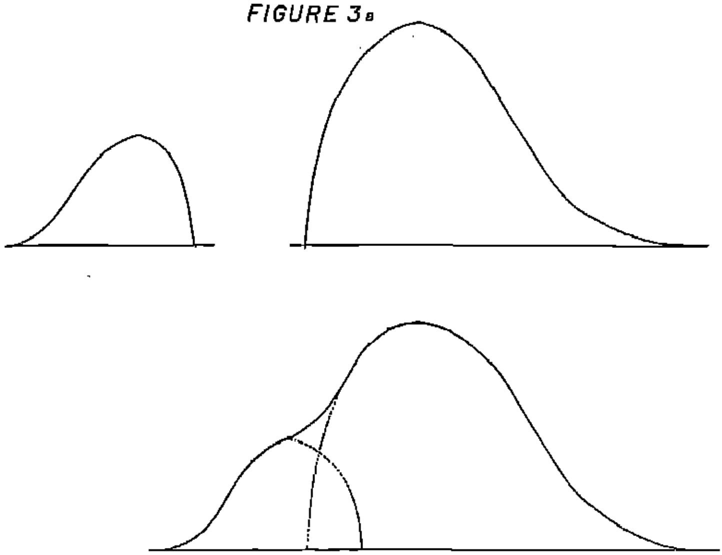
handedness based upon objective measures ought to yield more intelligible results if split into two distributions, the left-dominant and the right-dominant. The outer limits would, of course, be the extremes of lateral manual dominance, i.e., the extremely left-handed and the extremely right-handed. The inner limit of the distributions is set at zero-handedness, or the point at which there ceases to be any difference between the two hands. This inner limit in reality is not a point but a range, since the measurement error must be taken into account. The theory puts no limitations on the type of curve that may describe left-handedness or right-handedness. Such curves may vary in symmetry, range, and kurtosis. Figures 3A and 3B illustrate rather crudely what might be expected when two slightly overlapping distribution curves are combined into one distribution. In these illustrations the curve on the left half of the page is made approximately one-half the size of the curve on the right half of the page. This has been done to suggest the fact that the existence of two such separate distributions might be inconspicuous except in large samples if the smaller of the two distributions contained only a fraction of the cases in the larger distribution, as would be the case in distributions of handedness. Figure 3A shows the way two normal distributions might coalesce to give a distribution with a second mode. Figure 3B shows how two skew curves of Pearson's Type III might combine to give a curve only slightly suggestive of the fact that two distributions have been combined. The reasonableness of such combinations will be discussed later

FIGURE 3A



COALESCENCE OF NORMAL CURVES

FIGURE 3B



COALESENCE OF SKEW CURVES

when the data from the present investigation are considered.

Equivalence of Handedness Ratio

The remainder of this section will be concerned with the presentation and analysis of certain fundamental statistics, especially as they bear upon the problems of curve fitting. Table 7 presents some of the essential statistical constants for each of the measuring instruments used. As a prerequisite to a clearer understanding of what these values imply, it is necessary first to discuss in a more fundamental fashion the extent to which the handedness ratios used can be considered to be equivalent. Kelley (9, pp. 110-114) states the following three minimum conditions which must be met if ratios are to be considered as equivalent:

> 1. One point of the first scale must be known to be equal to one point of the second.
> 2. A second point of the first must be known to be equal to a second point of the second; and,
> 3. The law establishing the relationship between successive points on the first must be known to be the law underlying the second.

To what extent does the ratio $\frac{R-L}{R+L}$, or its variant form when errors are involved, $\frac{L-R}{R+L}$, meet these three requisite conditions? The first condition seems to be met in the case of the zero score. When the achievement or the lack of achievement, i.e., number of errors, of the right hand and the left hand are equal, the numerator of the fraction is zero, the ratio is zero,

TABLE 7

ESSENTIAL STATISTICS OF HANDEDNESS MEASURES

Statistic	Test: Target	Pin	Escape	Treasure	Composite	L-E-R
Median of total group	3.50	21.68	40.28	41.76	40.41	8.47
Mean of total group	4.95	21.50	38.39	39.42	35.42	7.00
Standard deviation of total group	6.55	14.65	24.00	28.70	19.52	4.32
Mean of right-handed	11.30	24.10	43.07	46.36	39.69	8.28
Standard deviation of right-handed	5.04	11.20	16.80	19.10	12.10	1.84
Mean of left-handed	—11.80	—19.15	—40.12	—42.71	—39.12	—6.78
Standard deviation of left-handed	3.38	8.95	16.90	20.60	12.10	2.41
Probable error of measurement	3.14	5.79	11.09	15.32	7.91	1.54
Percentage below probable error of measurement	4.76	4.34	4.34	4.27	4.91	6.54
Probable error of measurement range	—5 to +4	—5 to +4	—10 to +9	—10 to +9	—10 to +9	—2 to +2
Mean of total / S.D. of total	.756	1.468	1.599	1.374	1.814	
Mean of right-handed / S.D. of right-handed	2.242	2.152	2.564	2.427	3.280	
Mean of left-handed / S.D. of left-handed	3.491	2.140	2.774	2.073	3.233	
Number of cases	1281	1288	1244	1239	1244	2155

and the amount of handedness represented is zero-handedness or ambidexterity, and this is the same for all tests. The second condition is met, at least theoretically, in the 1.00 ratio, when a total lack of achievement in one hand is associated with optimum achievement in the other hand. In practice the calibration of the various tests is so different that a ratio of 1.00 is harder to achieve in some tests than in others. This factor alone, however, would not prevent the ratios from being considered as roughly equivalent. The third requirement, that the law establishing the relationship between successive points on the first must be known to be the law underlying the second, can be considered as met if the distribution curves describing the two variables are similar in type and if the ratio of the measure of variability to the measure of central tendency is the same for each test. It is at this point that the ratios used in this study lose in comparability. The curves describing the various distributions are not all similar and the ratio of the mean to the standard deviation is not the same for all tests. However, the distribution curves have many characteristics in common and the ratios of the mean to the standard deviation are sufficiently similar to say safely that, if the comparison is not pressed too hard, a given ratio on one test is roughly comparable to the same size of ratio on another test with the possible exception of the Target Test. These ratios of the mean to the standard deviation are given in Table 7, not only for the total distributions covering the entire range from extreme left-handedness to extreme right-handedness, but also for the two

tails of these larger distributions representing right-handedness on the one hand, and left-handedness on the other. If the hypothesis outlined earlier in this chapter, which suggests that the distributions of handedness are better considered as separate entities, is sound, this is a logical step and the results seem to be entirely reasonable; in fact, when the ratios of the mean to the standard deviations for the distributions of right-handed ratios are considered, the handedness ratios obtained with the Target Test would seem to be much more nearly comparable to the other test ratios than is indicated by the ratio of the mean to the S.D. for the total distributions.

Differences in Average Amounts of Handedness

The importance of this discussion of the comparability of the handedness ratios from test to test arises from the fact of the wide differences found in the means and standard deviations for the various tests. If the separate ratio scores are comparable, the statistical constants are even more likely to be comparable. Therefore, it would seem to be entirely safe to say that there are real differences in the degree of average handedness shown and in the variability of handedness from test to test—differences which must be due to fundamental variations in the traits measured. The increase in the means and the increase in variability seem to be coincidental with the involvement of progressively higher coordinations. This conclusion is strengthened when we note that the progression is op-

posite in direction for the distributions of handedness considered separately; that is, left-handedness, as well as right-handedness, becomes more pronounced as the tests involve finer and finer coordinations.

The No-Handedness Range

In obtaining the separate distributions of right- and left-handedness mentioned above, the point or range separating right-handedness from left-handedness had to be decided upon. The necessity for doing this has already been discussed (footnote 5), along with the method finally adopted, namely, the application of the probable error of measurement to the zero score. This was necessary because the unreliability of the tests used would not permit a sharp dichotomy at the zero point. Because of the grouping of scores into step intervals this criterion cannot be precisely applied, but that range of step intervals which nearest approaches the range of scores indicated by the probable error of measurement has been used. For example, in the Target Test, the P.E. of measurement is 3.14 and the most comparable range of step intervals is one step interval plus and minus from zero. And yet due to the bunching up of scores around zero this really does an injustice to the two separate handedness distributions. The percentage of scores below —5 is only 1.78 while the percentage of scores below—3 (the nearest whole number to 3.14) is 4.76. In the Treasure Test the P.E. of measurement is 15.32 and it was necessary to decide whether to use the step interval range from —10 to

+9 or the range from —20 to +19. It was felt that the wider range would cut off too many cases in view of the relatively small number of cases in the left-hand distribution, so the —10 to +9 range was chosen. The P.E. of measurement ranges for each test are shown in Table 7.

PERCENTAGE OF LEFT-HANDEDNESS

The percentage of left-handedness for each test of the battery, the Composite score, and the L-E-R Criterion are also given in Table 7. For the objective tests these percentages range from 4¼ to 5, while the L-E-R Criterion Questionnaire shows 6½% left-handed. While the difference between these percentages is not especially significant, there seems a slight tendency for those who feel themselves to be at least moderately left-handed to fall either in the ambidextrous range or even into the right-handed distribution when measured with objective tests.

TYPES OF CURVES FOUND

Figures 4 to 9 show the graphed distributions for each variable and Table 8 gives the essential information necessary for the determination of curve type.

A cursory glance at the Betas makes it clear that in no case is anything like a normal curve found, i.e., β_1 equal to 0, and β_2 equal to 3. The square root of β_1 is sometimes used as a measure of skewness and this value is given. From these values we see that every curve is skewed, and all but the Target Test are skewed

FIGURE 4

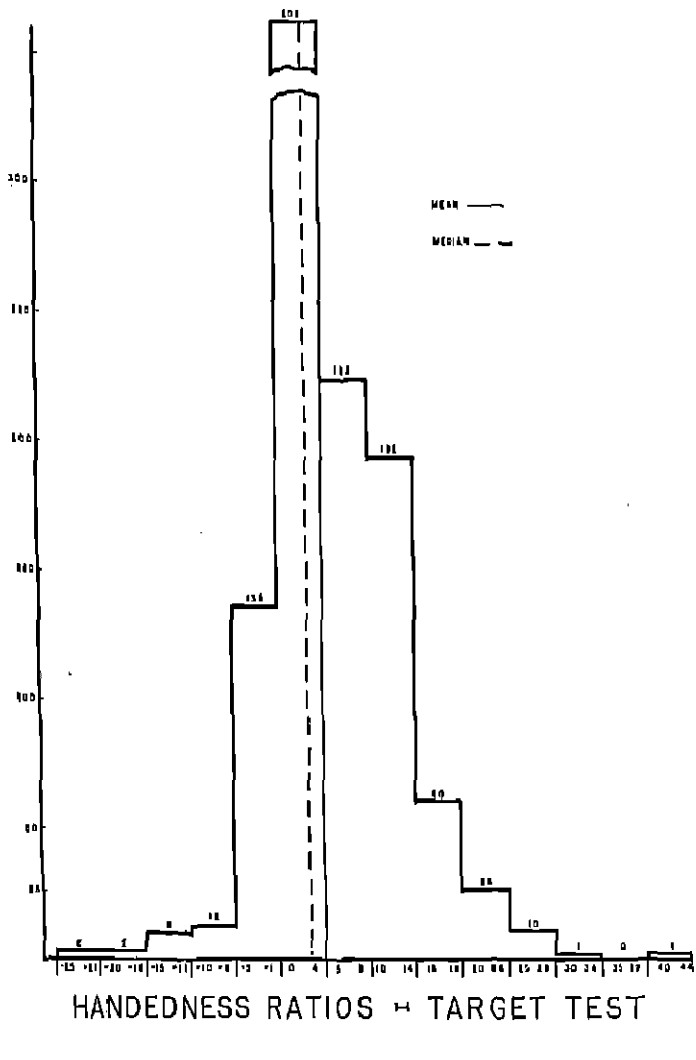


FIGURE 5

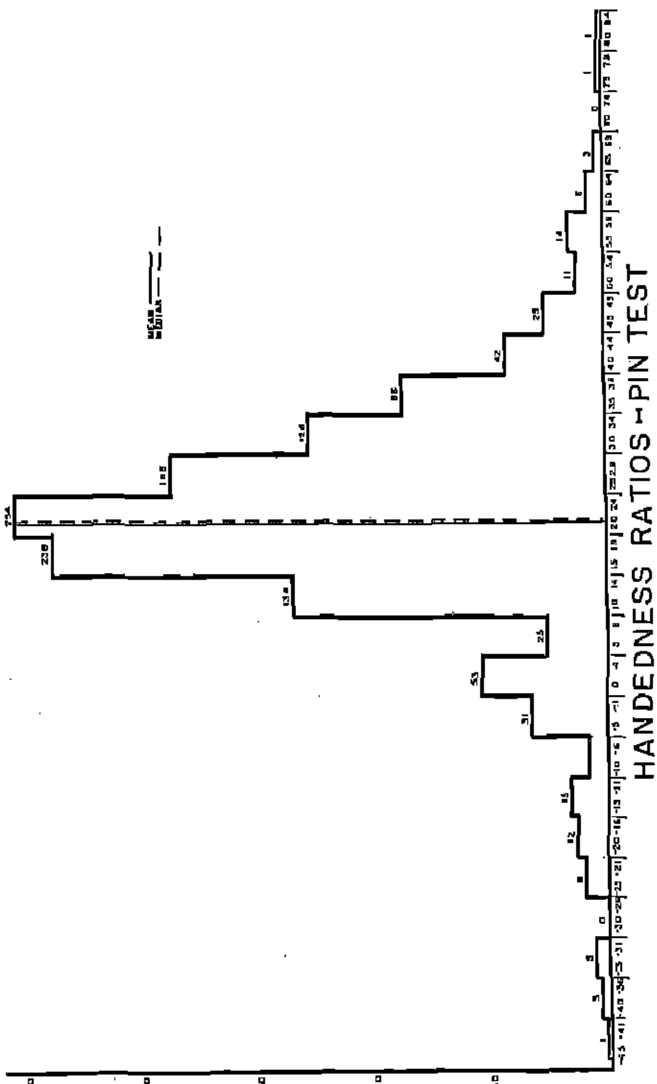
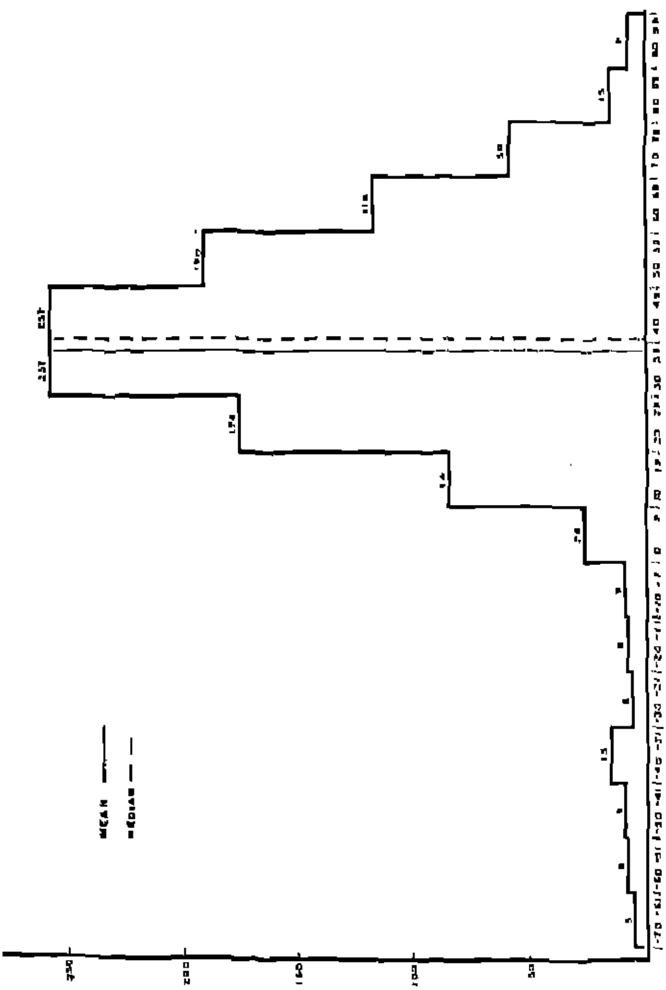


FIGURE 6



HANDEDNESS RATIOS—ESCAPE TEST

FIGURE 7

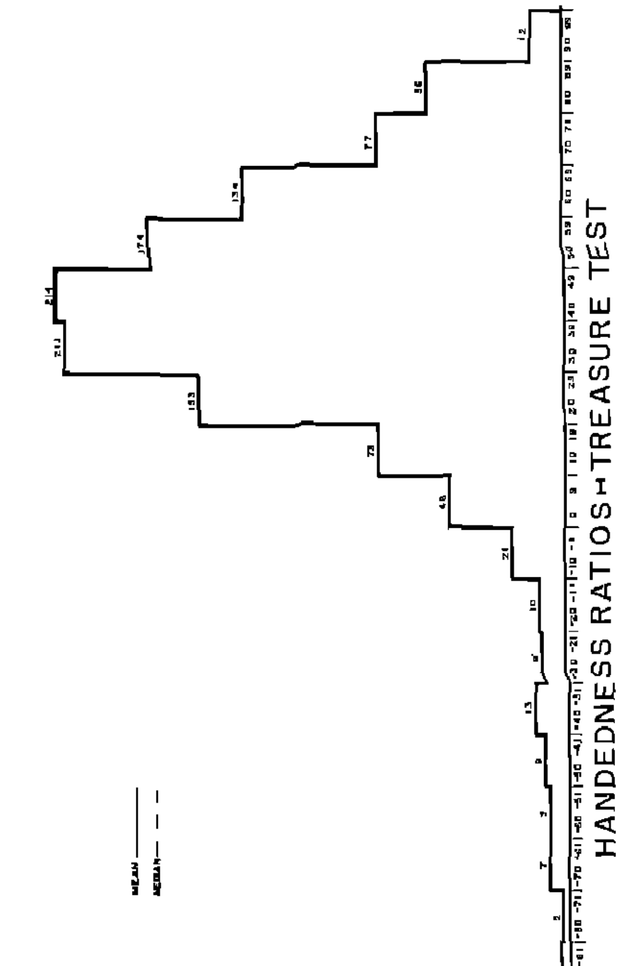
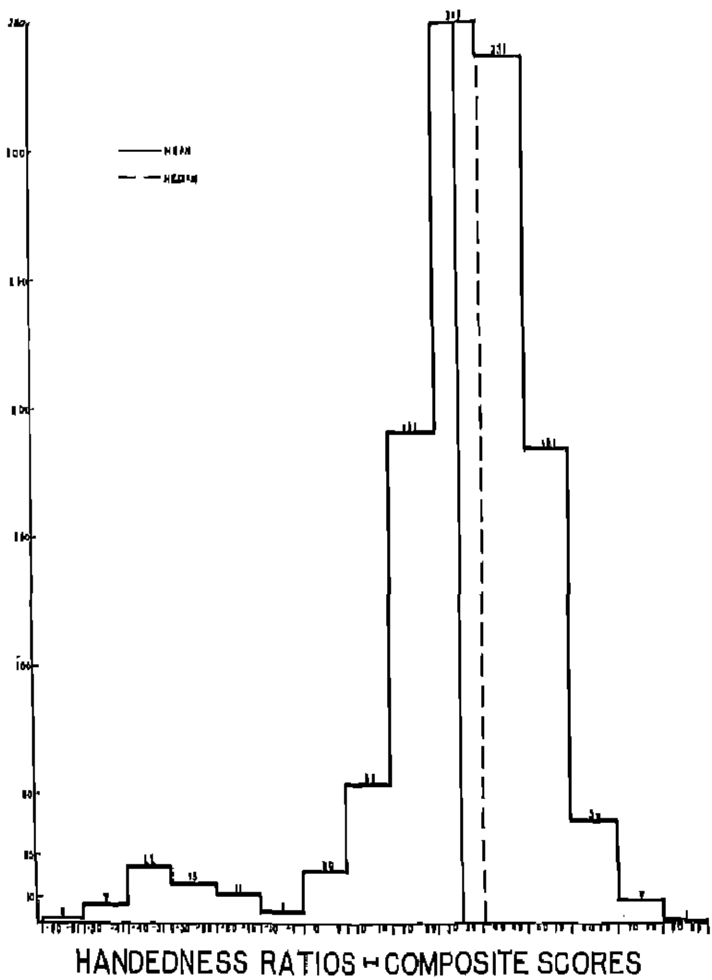


FIGURE 8



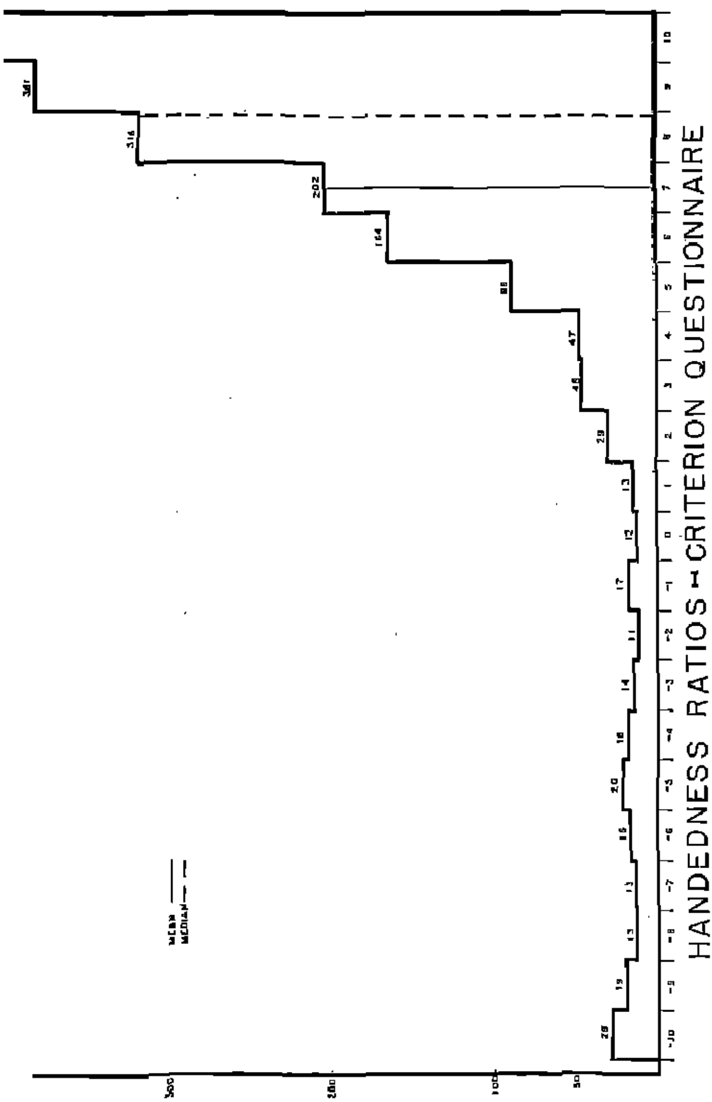

HANDEDNESS RATIOS - CRITERION QUESTIONNAIRE
MEAN
MEDIAN
316
202
164
88
47
300
200
100
50

TABLE 8

CRITERIA FOR DETERMINING FREQUENCY TYPES

Test	β_1	β_2	K_1	K_2	Type of curve	$Sk=\sqrt{\beta_1}$	$Ku=\beta_2-3$
Treasure	1.370	5.740	1.370	1.013	V	1.70	2.740
Escape	1.678	6.558	2.081	.869	V	1.295	3.558
Pin	.156	5.264	4.060	.032	Heterotypic	.395	2.264
Target	.579	5.464	3.189	.162	Heterotypic	.761	2.464
Composite	3.351	7.940	— .173	—26.642	III	1.831	4.940
L-E-R	5.370	8.092	—5.925	— 1.715	I_u	2.317	5.092

negatively in the sense that the median is numerically larger than the mean. All skewness values are statistically significant but this is due more to the large number of cases which reduces the S.E. than to the size of the skewness values.

For a normal curve, β_2 equals 3; therefore $\beta_2—3$ is often used as a measure of kurtosis. Negative values indicate that the distributions are platykurtic and positive values indicate that they are leptokurtic. All values of $\beta_2—3$ for these distributions are positive and large in relation to the norm of three, showing that the distributions are decidedly more peaked than a normal curve.

An attempt has been made to determine the curve type for each of the curves of the test battery and for the composite score. The entry under Curve Type in Table 8 has been determined in two ways: first, by finding the type of curve suggested for given values of K_2 in Pearson's *Tables for Statisticians* (11, see p. lxiii of the Introduction) and, second, by plotting a point on Diagram XXXV, page 66, the values of the *x* and *y* ordinates of this diagram being β_1 and β_2 respectively. The probable errors of the Betas and the Kappas of Table 8 are not given because with one exception either β_1 or β_2 or both were too large for these values to be determinable from Pearson's Tables. However, it seems clear from the tables that these P.E.'s would be relatively large, indicating that the distributions obtained are not very stable. Only one incontrovertible fact seems to come from these data, namely, that the type of curve which describes handed-

ness is not to be determined by any one test alone, if there is any one curve that can be said to describe it. For two distributions the data would seem to be satisfied by a Type-V curve (Treasure and Escape Tests); two others, the Pin and the Target distributions, are heterotypic, and for the Composite Score, if a value of —.1734 is considered near enough to zero to satisfy the equation $2\beta_2—3\beta_1—6$ equals 0 (a value of zero would seem definitely to be within the P.E. of the value —.1734), it would seem that a Type-III curve would satisfy these data. The next logical step would seem to be to fit a theoretical curve by the method of Least Squares, at least to the distribution of Composite Scores. However, the writer feels that the labor involved in doing this would not be profitably spent until some further work has been done to evaluate the suggestion made earlier in this section that these obtained distributions for each test might more adequately be considered as two separate distributions merged into one. This will involve fitting curves from the tails of the distribution where the data may be presumed to be more homogeneous and it is felt that the data at hand for the left-hand end of the distribution is not entirely adequate for this purpose. Therefore, since this possibility was not considered by the writer prior to the collection and analysis of these data, he feels justified in postponing this matter with the definite intention of carrying out such additional analysis in the reasonably near future.

THE L-E-R QUESTIONNAIRE DISTRIBUTION

The data concerning the distribution of scores ob-

tained with the L-E-R Questionnaire have not been discussed up to now, although presented in Tables 7 and 8. These data are not felt to be an integral part of the discussion preceding and therefore will be considered separately. These scores do not represent the measurement of an actual inherent skill or quality, but represent an expression of opinion concerning certain common habitual acts or performances. (See IV.) The rough gradation into "right," "either," or "left" is the only differentiation that can be made between the hands for any given question and the qualitative value of the questions is not the same. For example, many persons would say they could drink with the glass in either hand if the circumstances demanded it although they would prefer to use their dominant hand, but relatively few people could truthfully say that they write with either hand with anywhere near the same ease; therefore any such unweighted scale yields relatively little information of value concerning the real or inherent differences between the hands. The U-shaped distribution obtained is much as one would expect. The surprising element is the slow rate of rise for the left-hand group. This presents visible evidence of the degree to which left-handed people succumb to the environmental pressure to use their right hands, while the use of the right hand for right-handed people is reinforced by these pressures.

Summary

By way of summary, it may be useful to reconsider the four suggestions or theories presented earlier in

this section in the light of the evidence presented from this investigation.

1. Handedness is best described as a two-category trait or, as a modification of this, the distribution would logically be expected to be U-shaped.

The only evidence to bear out this theory is the shape of the distribution curve obtained with the L-E-R Criterion Questionnaire. If the questionnaire were longer, if proper weights could be determined for each question, and if finer gradations of handedness could be reliably determined, it is very doubtful if the U-shape of the distribution would persist. Furthermore, such a distribution would still be largely determined by strength of preference, a function of environment, and not by quality inherent in the make-up of the individual.

2. Handedness is best described by some I-shaped distribution of unknown skewness and kurtosis with the normal curve as a special case within this hypothesis.

The data of this investigation show no tendencies toward normality if the whole range of handedness is considered as one distribution; however, all curves are I-shaped curves, varying in skewness and kurtosis according to the degree of small muscle coordination and nervous control required. The distributions obtained in this investigation are atypical, however, as regards the relatively large size of the Betas obtained.

3. Handedness is truly bimodally distributed. There is evidence in at least three of the distributions obtained of a second mode which might logically be

thought of as suggesting the central tendency of the left-hand distribution. The trouble with this hypothesis is the difficulty in demonstrating that the data are truly homogeneous.

4. Handedness could more logically be considered as having two distributions, that of the left-dominant and that of the right-dominant, with curves similar in type but converse in skewness when not symmetrical. The data of this investigation are not inconsistent with this hypothesis. The persistent popular dichotomy of right-handed people and left-handed people and the equally persistent assumption that the two are somehow *different in kind seems justified.* We may think of people as being less and less right-handed only so long as the right hand maintains its advantage over the left. When the left takes the ascendency the logical demands of the situation are to consider such persons as left-handed.

If a majority of people on a majority of tests, objective in nature and not specifically calling for skill by reason of training, were found to be within the ambidextrous range, it might be adduced that the natural state of man is ambidexterity and the true distribution of handedness would consist of plus and minus variations from zero-handedness as a mean. There are, however, no indications that this is true.

However, the fact that such a distribution with the mode within the ambidextrous range is obtained when *a test involving the larger muscle coordinations is used* (Target Test) suggests the possibility that the progress of evolution has been in the direction of specialization of function and away from ambidexterity.

VI

HANDEDNESS RELATED TO CHRONOLOGICAL AGE, INTELLIGENCE, SCHOOL ACHIEVEMENT, AND CONDUCT FOR CHILDREN 9-15 YEARS OF AGE

Change in Handedness with Change in Age

Does handedness change in any systematic fashion from year to year as the child grows older? In other words do people grow more nearly ambidextrous as they grow older or do they become more than ever set in the direction and quality of their hand dominance? The evidence from this investigation on this point is inadequate but very suggestive. These data are presented in Tables 9 and 10. If handedness is thought of as a continuous homogeneous variate ranging from one extreme of dominance to the other, the proper *statistic to use in studying the variations in handedness* from year to year would be the mean of the total distribution. Therefore Table 9 presents these statistics for each year level from 9 to 15 inclusive and for each test of the battery and the composite score. An examination of the section of the table giving the means of the entire distribution reveals definite trends in three of the four tests. In the fourth, the Pin Test, the variation in the means is inconsistent in direction. For the Target Test the trend is definitely downward in the direction of ambidexterity while in the case of the Treasure and Escape Tests the trend is quite as defi-

TABLE 9

ANALYSIS OF GROWTH TRENDS FOR SEVEN C.A. LEVELS ON EACH TEST OF THE BATTERY AND FOR THE COMPOSITE SCORE

	Total group				Right-handed			
	N	M	S.D.	SE_M	N	M	S.D.	$S.E._M$
Target Test								
9 year olds*	80	5.69	5.29	.59	52	8.73	2.56	.36
10 " "	165	8.00	6.88	.54	116	11.31	2.48	.23
11 " "	189	7.05	6.97	.51	128	10.67	4.99	.44
12 " "	213	5.31	6.65	.47	111	10.24	5.10	.48
13 " "	237	5.04	6.08	.39	127	9.32	4.39	.39
14 " "	166	4.47	5.14	.40	77	8.88	3.66	.42
15 " "	84	2.36	4.75	.52	23	8.09	1.89	.39
All ages	1134	5.62	6.46	.19	634	9.97	4.83	.19
Pin Test								
9 year olds	79	24.66	16.18	1.82	73	27.96	11.63	1.36
10 " "	163	23.35	14.44	1.13	153	25.50	11.46	.93
11 " "	186	21.97	14.66	1.07	173	24.77	10.54	.80
12 " "	215	19.88	15.46	1.06	186	24.20	11.56	.85
13 " "	239	20.66	14.24	.92	217	23.31	12.97	.88
14 " "	175	22.57	14.27	1.08	163	24.67	12.21	.96
15 " "	77	18.62	14.90	1.70	69	22.00	10.14	1.22
All ages	1132	21.45	14.65	.44	1032	24.49	10.83	.34
Treasure Test								
9 year olds	77	25.92	21.91	2.50	65	35.73	15.18	1.88
10 " "	166	30.64	24.67	1.91	143	37.79	16.19	1.35
11 " "	175	32.99	24.21	1.83	162	40.55	17.51	1.38
12 " "	208	39.36	30.17	2.09	184	47.87	17.91	1.32
13 " "	233	42.61	27.61	1.81	213	49.01	16.99	1.16
14 " "	168	43.96	25.13	1.94	156	48.54	18.77	1.50
15 " "	87	46.68	27.06	2.90	82	51.08	19.73	2.18
All ages	1113	38.87	26.77	.80	1005	45.08	18.05	.57
Escape Test								
9 year olds	74	29.09	19.42	2.26	67	33.90	11.99	1.46
10 " "	165	34.32	20.75	1.62	156	37.64	23.57	1.89
11 " "	171	35.17	22.20	1.70	153	40.84	15.13	1.22
12 " "	206	35.33	28.97	2.02	184	42.98	16.79	1.24
13 " "	238	39.70	25.48	1.65	223	44.72	16.76	1.12
14 " "	178	43.26	25.37	1.90	165	49.04	11.46	.89
15 " "	82	41.08	21.84	2.41	80	43.25	16.87	1.89
All ages	1114	37.36	24.81	.74	1028	42.73	15.07	.47
Composite Score								
9 year olds	75	33.43	17.72	2.04	71	36.61	11.62	1.38
10 " "	163	36.16	17.09	1.34	154	39.11	12.07	.97
11 " "	176	36.60	18.49	1.39	164	40.35	11.97	.92
12 " "	208	34.36	21.84	1.51	190	39.66	12.62	.92
13 " "	232	36.74	20.20	1.33	218	40.83	11.72	.79
14 " "	173	35.02	21.32	1.62	159	40.04	12.38	.98
15 " "	87	34.84	18.76	2.01	81	38.57	11.01	1.22
All ages	1113	35.59	19.70	.59	1052	39.43	12.26	.38

*Any level of Chronological Age is defined as the naming year plus and minus six months, e.g., the nine year olds include all those from 8 years and 7 months to 9 years and 6 months inclusive.

TABLE 10

THE SIGNIFICANCES OF THE DIFFERENCES BETWEEN THE MEANS OF THE 10-YEAR-OLD GROUP AND THE 14-YEAR-OLD GROUP FOR EACH TEST OF THE BATTERY AND FOR THE COMPOSITE SCORE

Test	Total group			Right-handed group		
	Diff.	$\sigma_{diff.}$	$\frac{\text{Diff.}}{\sigma_{diff.}}$	Diff.	$\sigma_{diff.}$	$\frac{\text{Diff.}}{\sigma_{diff.}}$
Target	— 3.53	.67	5.27	— 2.43	.84	2.89
Pin	— .78	1.56	.50	— .83	1.39	.60
Treasure	8.94	2.50	3.58	11.40	2.08	5.48
Escape	13.32	2.72	4.90	10.75	2.02	5.32
Composite	—1.14	2.11	.54	.93	1.39	.67

nitely upward, or in the direction of a wider differentiation between the hands as age increases. Table 10 gives the significance ratios for differences between the mean scores of the 10th and 14th years test by test. The difference between the 9th and the 15th years was not chosen because the number of cases was so much smaller for these two years than for the remaining five age levels studied. For the three tests just mentioned the significance ratios are all greater than 2.5, and therefore can be considered statistically significant. The direction of these differences is consistent with the trends mentioned above.

The means for the Composite Score show less fluctuation and the difference between the Mean Composite Score for the 9-year-old group and the Mean Composite Score for the 14-year-old group is not statistically significant.

If the range of handedness from extreme right-handedness to extreme left-handedness is considered as being best described not by one distribution but by two,

one of the right-handed and one of the left-handed, the proper statistics to use in studying growth changes would be the means of these two separate distributions in order to tell whether the trend in the means is consistently outward, i.e., in the direction of more extreme differences between the hands, or consistently inward, in the direction of ambidexterity. Unfortunately the number of cases in the left-handed distributions by age levels is not great enough to justify computing these means for the left-handedness distributions. However, the means for the right-handed group, i.e., all those having plus scores above the standard error measurement of the zero score, are given for the purpose of comparing them with the means of the entire distribution. The trends are consistent in both series of statistics. The Target Test shows a trend toward ambidexterity, and the Treasure and Escape Tests show the same trend toward a more confirmed right-handedness with increase in age. In the Pin Test the same random fluctuations are found with no consistency in direction. The difference between the means for the three tests where the trend is definite are all statistically significant.

These statistics take on meaning and relevance when they are considered in the light of the nature of the tests themselves. The Target Test involves the use of the larger muscles of the forearm and shoulder. Apparently as one grows older there is less differentiation between the arms so far as the coordination of the larger muscles is concerned. In the Pin Test, where the degree of coordination required is somewhat finer, no

definite trend is discernible; while in the Treasure and Escape Tests, where a high degree of coordination is necessary, the trend becomes definite in the direction of greater specialization. Of course in the latter case the influence of training is probably one of the factors responsible for the trend.

Handedness Related to Other Factors

The weight of the evidence available from other investigations concerning the relation between handedness and intelligence, school achievement, and conduct suggests that there is no significant difference between the left-handed group and the right-handed group in these particulars. The writer feels, however, that the evidence concerning these relationships is entirely too meager for any final judgment. Nor does he advance the evidence available from this investigation along these lines as being any more conclusive, because either the number of left-handed children is too small to justify any definite conclusions, or the measuring instrument used to measure the characteristic to which handedness is being related is of unknown reliability.

School Grades

School grades in work and conduct are the first measures with which handedness are related here. All available letter grades for each child were averaged by assigning the following values to the letters and getting the arithmetical mean: A=5, B+=4, B=3, C=2, D=1. Varying numbers of grades were available for different children ranging from one year's records

up to the records covering seven or eight years. Thus the reliability of the final estimate of work or conduct varies greatly from child to child, since it is obvious that the combined estimate of several teachers covering several years of work is more reliable than one teacher's estimate for one year. In selecting the right- and left-handed groups only those cases were included who had scores, either right- or left-handed, beyond the range of the standard error of measurement, applied to the zero score. The Composite Score is used as the best single measure of handedness and the approximate limits for the standard error of measurement applied to the zero score are —10 to +9. In other words, all those who had negative ratios below —10 on the composite score were included in the left-handed group and all those who had positive ratios greater than +9 are included in the right-handed group. In the writer's estimation, the cases included within the ambi*dextrous range would be the ones for which such com*parisons as are made here would be most interesting, since this range probably includes the majority of cases where there have been difficulties over enforced changes of handedness. However, the number of cases falling within this range in this study is altogether too small to make such comparisons of any value.

Table 11 gives the statistics concerning the relative average achievement of the right- and left-handed groups. It will be seen that the mean work grade and the mean conduct grade are both slightly higher for the right-handed group but that the differences are not statistically reliable.

TABLE 11

INTELLIGENCE, SCHOOL ACHIEVEMENT, AND CONDUCT OF LEFT- AND RIGHT-HANDED CHILDREN, ACCORDING TO THE COMPOSITE SCORE, COMPARED

	$N_{R.H.}$	$M_{R.H.}$	$\sigma_{R.H.}$	$\sigma_{M_{R.H.}}$	$N_{L.H.}$	$M_{L.H.}$	$\sigma_{L.H.}$	$\sigma_{M_{L.H.}}$	Diff.	$\sigma_{diff.}$	$\%_{L.H.}$
School work—all schools	1079	3.18	.74	.02	56	3.17	.73	.10	.01	.10	.048
School conduct—all schools	1083	4.10	.78	.02	56	4.04	.82	.11	.06	.11	.057
Otis-Class Test—P.S. 44											
Part I	278	40.27	17.90	1.07	17	39.94	19.85	4.82	.33	4.94	.056
Part II	278	35.00	14.55	.87	17	33.47	12.30	2.99	1.53	3.11	.056
National Intell. Test—											
P.S. 42 and 43	659	95.15	17.85	.70	29	91.83	18.40	3.42	3.32	3.49	.042

INTELLIGENCE

The Otis Classifications Test was given in one school and the mean raw scores for Part I and Part II are given in the table. Part I is the measure of educational achievement and Part II is the measure of intelligence. We see here that in both cases the differences are in favor of the right-handed group but are not statistically significant. The final comparison given in Table 11 is between N.I.T. Intelligence Quotients, which were available for two schools. For P.S. No. 42, from which approximately one-half of the cases were drawn, the tests were administered by the present investigator, for P.S. No. 43 the data were taken from school records and the date of testing is not uniform for all the children. There again the difference is in favor of the right-handed group and, while larger than the other differences, is not statistically reliable.

If these data have any significance at all it arises from the fact that the direction of the difference is consistently in favor of the right-handed group. The smaller number of scores in the left-handed groups with the resulting relatively large standard errors for the differences found gives plausibility to the hypothesis that larger distributions of left-handedness might yield differences which would be statistically reliable.

VII

SPECIFICITY OF HANDEDNESS

One of the persistently controversial issues in the measurement of manual laterality is the question of the specificity of hand dominance or hand preference. In discussing this specificity of handedness, the distinction should be made between hand dominance as measured by objective tests and hand preference measured by the proportion of habitual acts performed by the preferred hand. In measuring preferential handedness, any inventory of habitual performances will suggest that training plays a greater or less important part from act to act. As a result, handedness determined by any one act as a criterion will not necessarily be the same for even a majority of the other performances in the inventory. For example, many naturally left-handed people are made to write with the right hand; hence, if writing were taken as a criterion, many would be right-handed who would be found to be left-handed according to some other criterion such as throwing. Objective tests used to measure hand dominance minimize the influence of training, if those tests call for performances not specifically taught. In other words, training enters in only by transfer. Furthermore, such tests make it possible to measure all degrees of dominance from one extreme of laterality to the other, and make the statistical determination of co-variance more reliable.

TEST INTERCORRELATIONS

The obvious method to be followed in determining

the degree of association between tests of hand dominance is the correlational technique. The intercor-

TABLE 12

Reliability Coefficients and Intercorrelations for Handedness Test Battery, Composite Score, and L-E-R Criterion Questionnaire

		Target	Pin	Treasure	Escape	Composite	L-E-R
Target	Uncorrected	*.7024*	.3770	.1853	.2390	.6251	.3650
	Corrected	*.8252*	.4997	.2827	.3338	.8172	.4726
Pin	Uncorrected		*.8102*	.3288	.4185	.7649	.4711
	Corrected		*.8951*	.4672	.5444	.9311	.5679
Treasure	Uncorrected			*.6115*	.6255	.7486	.5249
	Corrected			*.7589*	.9364	1.0000	.7284
Escape	Uncorrected				*.7296*	.7712	.5736
	Corrected				*.8437*	.9892	.7288
Composite	Uncorrected					*.8330*	.6095
	Corrected					*.9089*	.7240
L-E-R	Uncorrected						*.8493*
	Corrected						*.9185*

relations found between the various tests of handedness used in this study are presented in Table 12. Before these data are discussed, however, some of the shortcomings of this method may be considered with profit. In the first place, the size of the correlation coefficient between any pair of tests is in part a function of the reliability of the tests, i.e., the less reliable the tests, the greater the degree of attenuation. To state this in another way, if the reliability coefficients of a pair of tests were very high and the intercorrelation between the tests were low, the assumption of a lack of relationship between the factors measured by the tests is more dependable than it would be if a low intercorrelation were accompanied

by low reliability coefficients. Even if the usual correction for attenuation is made, this is at best only an estimate of the degree of association which would be found if the tests used were more reliable.

On the other hand, the degree of association found between the tests may be due to the fact that the tests have a high degree of similarity. Therefore, it is important to answer the question: "Are the tests used in this investigation sufficiently different as to be clearly not variant forms of the same fundamental test pattern?" In the present case, the distinction between the tests seems to be clear in every case except the Treasure and Escape Tests which may very well be only variants of a fundamentally similar pattern. The Target Test is essentially a test of the ability to coordinate the musculature of the arms and shoulders. The Pin Test involves both speed of movement and coordination of the muscles of the fingers and hands. The Treasure and Escape Tests are essentially measures of steadiness with speed a non-essential factor. The L-E-R Criterion Score is obviously entirely different in derivation from the other tests of the battery. Thus it would seem that no violence is being done to the facts when it is assumed that at least four distinctly different approaches to the measurement of handedness are represented in this study. The use of the ratio of the difference between the scores to the sum of the scores of the two hands, as the measure of handedness, should have only the effect of discounting differences in the total achievement of different individuals on each of the tests while emphasizing the relative differences in the degrees of hand dominance.

Table 12 gives the reliabilities of each of the tests by the split-halves method, both the uncorrected, i.e., before the Spearman-Brown Prophecy Formula was applied to estimate the reliability of the total test from the correlation of the two halves, and the corrected coefficients.[6] It also gives the intercorrelations between the tests of the battery, the composite handedness score, and the L-E-R Questionnaire, both as raw correlations and when corrected for attenuation. All the intercorrelations are positive and the smallest coefficient, the correlation between the Target Test and the Treasure Test (.1853±.02), is more than nine times its probable error. No probable error for any of the coefficients given is as great as .03. The Target Test shows the least amount of correlation with the other tests but it does correlate positively with them all, even the L-E-R Questionnaire. From this low value of .1853 the coefficients range upward in size. Especially significant are the relatively high *r*'s between each of the tests and the composite score where the uncorrected coefficients run as high as .77; in other words, when the standard deviations of the various tests are made equal and the scores averaged, the correlation of this average or composite with each of the tests in turn is high and the coefficients are approximately of the same order as the reliability coefficients. The relatively high correlations between the objective tests of the battery and the subjective estimation of hand preference, represented by the L-E-R scores, is additional evidence to the effect that handedness as measured by

[6]See Section III for a more extensive discussion of reliability.

a variety of tests is definitely related; in other words, the evidence herein presented would indicate that there is a positive association between the degrees of hand dominance which an individual will show on a series of tests; and if tests with perfect reliability could be devised, the evidence of this association would, in all probability, be strengthened. The intercorrelations between tests of handedness, since they measure concomitant variation over the whole distribution from one extreme of handedness to the other, will not tell us what the chances are that a person who is left-handed according to one test will also be left-handed according to a series of other tests; in other words, knowing the degree of association, as measured by the correlation coefficient, will not tell us the degree of stability of handedness when we dichotomize our distribution into left-handed and right-handed. In order to throw some light upon this problem, we have sharply dichotomized our distributions at the zero point. All those below zero are considered as left-handed and all above zero as right-handed. The range of ambidexterity, used previously because of the errors of measurement, has been ignored in this case and the P.E.'s of the percentages are given as a means of estimating the degree of unreliability in proportions shown. Table 13 gives the number and percentage of the total distribution who are left-handed according to each test of the battery and the Composite Score. Using the number of left-handed, according to each test, as a base it gives the number and percentage who are also left-handed on each of the other tests and the Composite Score. For

TABLE 13

THE LEFT-HANDED FOR EACH TEST SELECTED BY EACH OTHER TEST USED AS A CRITERION

Test	Target		Pin		Test Treasure		Escape		Composite	
	N*	%	N	%	N	%	N	%	N	%
Target			148	.32 ±.03	148	.30 ±.03	151	.28 ±.03	141	.34 ±.03
Pin	76	.63 ±.04			78	.69 ±.03	76	.70 ±.03	75	.76 ±.03
Treasure	77	.58 ±.04	79	.68 ±.03			74	.73 ±.03	69	.75 ±.03
Escape	63	.67 ±.04	64	.83 ±.03	63	.86 ±.03			68	.89 ±.03
Composite	61	.79 ±.03	61	.93 ±.02	57	.91 ±.03	62	.94 ±.02		
Per cent below zero**		.12 ±.007		.07 ±.005		.06 ±.005		.06 ±.005		.05 ±.005

*N is given because N varies slightly from test to test, due to the fact that an occasional imperfect test has been dropped out.

**All negative scores are called left-handed in this case.

example, in the case of the Target Test, 148 are left-handed, or approximately 12% of the total distribution. Of these 148, 48 or 32% are also left-handed on the Pin Test. In the case of the Pin Test, 76 or 7% of the total distribution are left-handed; of these 76, 48 or 63% are also left-handed on the Target Test. It would not be profitable to rehearse the whole table in this manner, but it may be helpful to point out some salient facts. When the Target Test is associated with each of the others in turn, including the Composite Score, it is seen that the percentage of left-handed common to each pairing is in the neighborhood of 30%. When the other tests are associated with each other, we see that the percentage of left-handed common to both ranges from approximately 60% to approximately 90%, that is, twice and more as large in proportion as in the Target Test. This discrepancy is

probably explained by the fact that the Target Test isolates as left-handed approximately twice as many as any other test and two and a half times as many as are selected by the Composite Score as left-handed. Considering in the main the range of the percentages of left-handed cases common to any two tests, omitting the Target Test, (from 60 to 90%) and in particular the percentage of those left-handed according to the Composite Score who are also left-handed according to each of the other tests, i.e., 79% for the Target Test, 93% for the Pin Test, 91% for the Treasure Test, and 94% for the Escape, we must inevitably reach the conclusion that there exists a very substantial degree of consistency in the type of handedness shown from one test to another. Moreover, it would not be doing violence to the evidence now at hand to say that it seems quite justifiable to dichotomize the total population of *handedness into left-handed cases, on the one hand*, and into right-handed cases on the other hand. Such a statement evades the responsibility of dealing with the problem of ambidexterity. Probably the most reasonable assumption concerning ambidexterity is that a true measure of handedness sufficiently discriminatory to measure the smallest reliable difference between the hands would show practically no genuinely ambidextrous cases, although there might be a small percentage of cases where the discernible difference between the hands was of no practical importance, so far as the limitation it puts on the activity of the non-dominant hand is concerned.

VIII

SUMMARY, CONCLUSIONS, AND SUGGESTIONS FOR FURTHER STUDY

Main Purposes of the Investigation

The main purposes of the investigation were:

1. To develop a battery of reliable and valid group tests of hand dominance.
2. To apply this battery to a random sample of at least one thousand cases previously unselected as to handedness, and from this population to:
   a. Determine the percentage of left-handed.
   b. Determine the type or types of distribution curves for the various tests.
   c. Determine the specificity or association between handedness tests.
   d. Study changes in handedness coincident with changes in age.
   e. Determine, if possible, in what respects, if any, the left-handed are atypical.
   f. Develop a Criterion test which may be used to select additional cases of left-handedness, with a known degree of efficiency.

In the main these purposes have been achieved. The corrected reliability coefficients range from .76 to .91, values which compare favorably with the reliability coefficients for similar tests. The tests are as easy to give as the average intelligence test and are objective to score. The scoring, however, is very time consuming.

The tests were given to over 1300 children and from the analysis of this population the following conclusions are drawn:

1. *Percentage of Left-Handed.* The percentage of left-handed for each test of the battery and for the Composite Score ranges from 4¼ to 5% when the upper limits of left-handedness are determined by the probable error of measurement applied to the zero score. That is, all whose scores fall below the limits set by the probable error of measurement are considered as definitely left-handed. This seems a logical division since one whose score was on or near this point would be likely to have a score on the plus or right-handed side of zero by chance only 25 times out of one hundred measurements.

2. *Curve Types.* No one curve type describes all the test distributions. Rather, the shape of the distribution curve seems to be in part at least a function of the type of test used. As the tests require finer and finer coordinations the standard deviations increase, and the means of the separate distribution of the right- and left-handed move apart. All curves are skewed and the direction of the skewness is negative in all cases except the Target Test. As the tests require finer and finer coordinations the negative skewness becomes more marked and the left-handed end of the distribution takes on more of the characteristics of an independent distribution. All distributions are leptokurtic. Perhaps the most important conclusion concerning the distributions of handedness is that they show no evidence of being normal curves.

3. *Specificity of Handedness.* All handedness measures are positively correlated with the r's ranging from .2827 to 1.00 when corrected for attenuation. The percentage of overlapping for successive pairings of the tests when each test is taken in turn as the criterion of left-handedness, ranges from 32% to 94% with the majority of the values above 50%. From these data it seems safe to conclude that handedness, as measured by objective tests of this nature, is essentially unitary and not specific to the type of test. Later, some method of factor analysis will be applied to these data and the results will be reported in a separate article.

4. *Handedness and Chronological Age.* Handedness changes, coincidental with chronological age, were found to be relatively slight. For the Pin Test the changes were inconsistent. For the Treasure and Escape Tests they were consistent in the direction of more decided handedness for the upper age levels. The difference between the means of the 10- and 14-year-old groups was fairly large and was statistically significant. For the Target Test the change was opposite in direction, that is, tending toward ambidexterity. In this case as well, the difference between the 10- and 14-year-old groups was large enough to be statistically significant. For the Composite Score the changes were inconsistent from level to level, reflecting the opposite influence of the Target *vs.* the Treasure and Escape Tests and the difference between the 10th- and 14th-year levels was not significant. From these data we may conclude that changes in handedness with chrono-

logical age depend upon the nature of these tests under consideration; when the tests involve relatively gross coordinations, the changes are in the direction of equalizing the difference between the hands as the child grows older, and when the tests demand a relatively high degree of coordination and muscular control, the changes with age are toward greater differentiation between the hands.

5. *Handedness Related to Achievement in Other Directions.* All differences found between the left-handed and the right-handed in school achievement in work and conduct, and in achievement on subject matter and intelligence tests, favor the right-handed group, but no differences were found to be statistically significant. Because of the relatively small number of left-handed cases, the investigation is inconclusive on these issues and any more definite findings must wait upon the accumulation of more data.

6. *The Criterion Questionnaire.* The attempt to develop a criterion to be used to segregate more left-handed cases for further study was comparatively successful considering only the data having to do with the Composite Score, 142 out of approximately 1300 children having L-E-R scores between —10 and +5. This includes 95% of those who were definitely left-handed according to their Composite Scores with a 5% loss, and 92% of those who either were left-handed or ambidextrous according to their Composite Scores with 8% loss. Only 25% of the 142 were left-handed while 32% were either left-handed or ambidextrous. In other words, when the criterion is used, 75% of those

selected will not be left-handed and 68% will be neither left-handed nor ambidextrous. These last two percentages may be called the percentage of inefficiency, which is noticeably reduced if the narrower range of Criterion scores from —10 to —1 is taken, but the percentage of efficiency is also reduced to 87% for the left-handed and 82% for the left-handed and ambidextrous together.

Recommendations for Further Research

This investigation ought to be used as the starting point for a series of further studies along a number of different lines. A few such possibilities for further research are suggested here.

Using the Criterion Questionnaire, left-handed and ambidextrous cases should be selected and measured with the test battery. Using these additional left-handed cases as the main experimental population, further deduction could then be made concerning the ways in which the left-handed do significantly deviate from the right-handed, if such deviations exist at all.

The ambidextrous group could also be studied as a class to see if they deviate significantly from either the left-handed or the right-handed.

If the larger sample of left-handed cases were comparable in other respects with the sample of right-handed cases secured in this study, the distribution of scores for the left-handed and for the right-handed could be compared and some more reliable conclusions reached concerning the hypothesis that handedness is

better described by two distributions than by one continuous curve taking in all degrees of handedness from one extreme to the other.

If the handedness battery could be given to a selected group who also were tested by some of the indirect methods of measuring handedness, such as the method of action current precedence, the relation between these two approaches to the problems of laterality could be established. If more adequate tests of eyedness which would yield a continuous distribution of the degrees of eye dominance could be devised, these tests might also be related to handedness as measured by the test battery.

The test battery or some modification of it might be used in Child Guidance Clinics to measure the direction and degree of hand dominance for clinic cases. If some method could be devised for collating the results of the use of these tests in wide-spread clinic situations with the other psychometric tests also employed clinically, it might be possible to throw some light upon the persistent and yet contradictory reports concerning the effect of handedness and especially the change of hand dominance upon the emotional stability and life adjustments of the child. The relation between handedness and stuttering could be established in a similar fashion.

References

1. Anastasi, A. The influence of practice upon test reliability. *J. Educ. Psychol.*, 1934, **25**, 312-335.
2. Chamberlain, H. D. The inheritance of left-handedness. *J. Hered.*, 1928, **19**, 557-559.
3. Dunlap, J. W., & Kurtz, A. K. Handbook of statistical nomographs, tables, and formulas. Yonkers, N. Y.: World Book Co., 1932. Pp. vii+163.
4. Fisher, R. A. Statistical methods for research workers. (4th ed.) London: Oliver & Boyd, 1932. Pp. xii+307.
5. Gesell, A. Infancy and human growth. New York: Macmillan, 1928. Pp. xvi+418.
6. Haefner, R. D. The educational significance of left-handedness. *Teach. Coll. Contrib. Educ.*, 1929, No. 360. Pp. vii+84.
7. Heinlein, J. H. A study of dextrality in children. *J. Genet. Psychol.*, 1929, **36**, 91-119.
8. Jones, M. C. The development of early behavior patterns in young children. *Ped. Sem.*, 1926, **33**, 537-585.
9. Kelley, T. L. Statistical method. (2nd ed.) New York: Macmillan, 1924. Pp. xi+390.
10. Koch, H. L., *et al.* A study of the nature, measurement, and determination of hand preference. *Genet. Psychol. Monog.*, 1933, **13**, 117-221.
11. Pearson, K. Tables for statisticians and biometricians. Part I (2nd ed.) London: Cambridge Univ. Press, 1924. Pp. lxxxiii+143.
12. Wellman, B. L. The development of motor coordinations in young children. *Univ. Iowa Stud.: Stud. Child Welfare*, 1926, **3**. Pp. 93.

APPENDIX

Nomograph for the Computation of Handedness Ratios

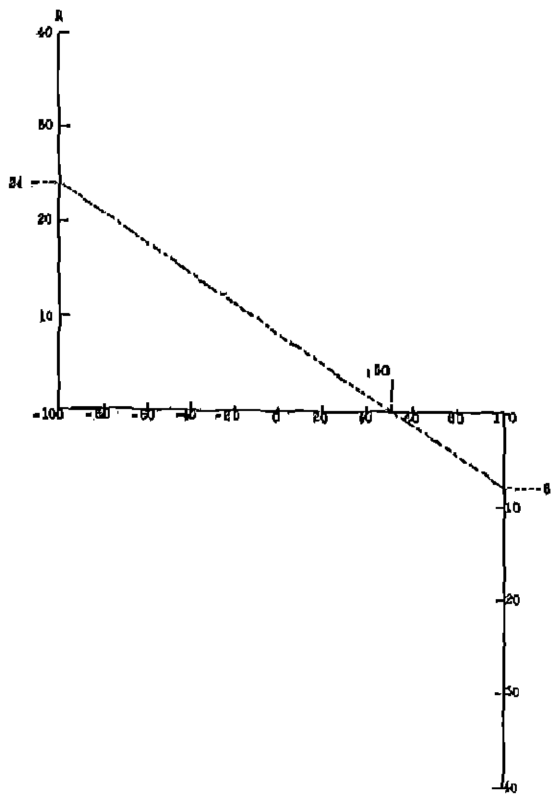
$$\text{Formula } \frac{R-L}{R+L} = r$$

R is the score of the right hand obtained on any given test.

L is the score obtained with the left hand.

r is the Handedness Ratio or the proportion that the difference *between the two hands is to the total achievement of both hands* together.

Data needed to enter the nomograph—scores of right and left hands.



Directions: Find the R and L values on the R and L scales. Draw a line or use a hair line drawn on a celluloid strip to connect the two points exactly. The desired ratio will be found on the horizontal scale. If R is greater than L, the ratio will be positive and right-handedness will be indicated. If R is less than L, the ratio will be negative and left-handedness will be indicated.

Example: The dotted line drawn on the illustrative nomograph connects the scores R=24 and L=8. R is greater than L and the ratio, therefore, will be positive. The point at which the dotted line intersects the horizontal scale will indicate the degree of handedness which, in this case, is +.50.

The nomograph given here is for illustrative purposes only. In actual practice, the nomograph used was drawn on standard 10 x 10 Keufel & Esser cross-section paper and the R and L scales were run up to 350. A line scratched on a celluloid strip was used as a hair line. The nomograph gave results accurate to the second decimal place, which was sufficiently accurate for all practical purposes since the scores were grouped in the statistical analysis with a step interval of 5 as the smallest unit. When the scores were less than 35 for each hand as sometimes happened in the Treasure and Escape Tests, the zeros were disregarded on the 350 scale and the larger units were used with a considerable gain in accuracy, i.e., 250 was considered as 25 or 240 as 24.

This nomograph is fundamentally similar to Nomograph No. 59 Percentages, given by Dunlap and Kurtz (3).

LE DÉVELOPPEMENT D'UNE BATTERIE DE TESTS OBJECTIFS COLLECTIFS DE LA LATÉRALITÉ MANUELLE AVEC LES RÉSULTATS DE LEUR APPLICATION À 1300 ENFANTS

(Résumé)

La mesure de la latéralité manuelle au moyen de tests objectifs tels que les tests de tapping et de stabilité a été limité pour la plupart au laboratoire psychologique. Cette etude rapporte le développement d'une batterie de quatre tests objectifs collectifs comparables à bien des égards au type du laboratoire. Le test de cible dans cette batterie emploi une petite lance, un bois rond long de 10 pouces avec un goujon à pointe d'épingle mis dans une extrémité avec un noeud élastique en forme de noeud coulant attaché à l'autre. On fait glisser la main à travers le noeud qu'on serre autour du bras de sorte que le poignet est bien serré et qu'il faut employer les muscles du coude, du bras, et de l'épaule. On a permis trois lancements dirigés vers chacune de soixante cibles. Dans le test de pointillage ou "tapping," on s'est servi d'une épingle à tête de verre pour piquer chaque cercle du test aussi vite que possible. C'est un test de temps lequel mesure et la vitesse et la précision. Deux tests de stabilité sont basés sur une série de figures concentriques, dont les côtés forment des voies d'une largeur décroissante à travers lesquelles on tire une ligne au crayon en essayant de ne pas toucher les côtés de la figure.

Les coefficients de constance selon la méthode "split halves" basée sur environ 650 cas, quand corrigés varient de 0,76 à 0,90.

On n'a pas établi la validité selon la méthode des corrélations à cause du manque d'un critère satisfaisant mais l'analyse logique des tests n'a révélé aucun vrai préjugé en faveur de l'une ou de l'autre main.

Le résultat pour l'usage des mains a été la proportion entre la différence dans le rendement des deux mains et le rendement total des mains, c'est-à-dire, $\frac{R-L}{R+L}$, la variation théorique des résultats étant ainsi de —1,00 à +1,00.

Les résultats des tests montrent que les tests diffèrent beaucoup dans le degré moyen de l'usage des mains et dans la variabilité, mais entre quatre et cinq pour cent du groupe entier sont gauchers dans tous les tests. Les courbes de distribution basées sur 1300 cas ne sont pas entièrement semblables dans tous les tests mais les tests où il s'agit des plus grandes quantités de coordination montrent des témoignages définis de la bi-modalité, dont la nature n'est pas claire d'après ces données à cause du petit nombre de cas dans le segment gauche dominant de la courbe. On n'a trouvé aucune relation non équivoque entre l'usage des mains et l'âge chronologique ou le rendement scolaire dans ces données.

On a testé un questionnaire de dix questions, opposé à cette batterie de tests, dans le but de choisir d'autres cas de gaucherie et il s'est montré d'une valeur de plus de 90% dans la sélection des gauchers selon une énumération composée des résultats où tous les quatre tests ont été considérés de valeur égale.

DUROST

DIE AUFSTELLUNG EINER BATTERIE VON OBJEKTIVEN GRUPPENSTESTS DER HANDVORHERRSCHUNG ZUSAMMEN *MIT DEN ERGEBNISSEN IHRER ANWENDUNG* BEI 1300 KINDERN

(Referat)

Die Messung der Handvorherrschung durch objektive Tests wie Klopfen und Stetigkeitstests hat sich zum grossen Teil auf das psychologische Laboratorium beschränkt. Dieses Studium berichtet über die Aufstellung einer Batterie von vier objektiven Gruppentests, welche in mancher Hinsicht dem Laboratoriumtypus vergleichbar sind. Bei dem Zielscheibentest in dieser Batterie wird ein Zielscheibenwurfspeer verwandt, der aus einem 10 Zoll langen hölzernen Pfeil besteht, in den ein gespitzter Dübel an einem Ende eingesteckt und mit einer elastischen Schlinge am anderen Ende gebunden wurde. Die Hand wird durch die Schlinge geschlüpft, die um den Arm festgezogen wurde, damit das Handgelenk kräftig befestigt und der Gebrauch des Ellenbogens, Arm- und Schultermuskel erzwungen wurde. Drei Würfe auf jede von sechsig Zielscheiben wurden gestattet. Bei dem Tüpfeln- oder Klopfentest wurde eine glasköpfige Stecknadel möglichst schnell zum Durchstechen jedes Kreises in dem Test gebraucht. Dies ist ein Zeittest, der sowohl Schnelligkeit wie Genauigkeit misst. Zwei Stetigkeitstests bestehen aus einer Reihe von konzentrischen Figuren, deren Seiten Wege von abnehmender Breite bilden, durch die ein Strich mit einem Bleistift gezogen wird, während der Versuch gemacht wird, die Seiten der Figur nicht zu berühren.

Die Zuverlässigkeitskoeffizienten vermittelst der gespalteten Hälftenmethode (split-halves method), welche ungefähr 650 Fälle umfassen, rangierten von 0,76 bis 0,90, wenn sie verbessert wurden.

Die Gültigkeit wurde durch die Korrelationsmethode deswegen nicht festgestellt, weil es kein angemessenes Kriterion gab, aber die logische Untersuchung der Tests wies keinen wesentlichen Vorteil zugunsten einer der beiden Hände auf.

Der Handvorherrschungswert war das Verhältnis zwischen dem Unterschied in der Ausführung der beiden Hände und der Gesamtausführung der Hände, d.h., $\frac{R-L}{R+L}$, woraus folgt, dass der theoretische Umfang der Werte sich von —1,00 bis +1,00 erstreckt.

Die Ergebnisse zeigen, dass die Tests sehr in dem Durchschnittsgrad der Handvorherrschung und in der Veränderlichkeit abweichen, aber zwischen vier und fünf Prozent der Gesamtanzahl der Vpn. sind linkändig bei allen Tests. Die Verteilungskurven von 1300 Fällen sind einander von Test zu Test nicht ganz ähnlich, aber diejenigen Tests, die das grösste Zusammenwirken verlangen, zeigen bestimmte Zeichen der Bimodalität, deren Natur aus diesen Daten nicht klar ist, wegen der kleinen Anzahl der Fälle in dem linksherrschenden Segment der Kurve. Kein unzweideutiges Verhältnis zwischen Handvorherrschung und chronologischem Alter, Intelligenz oder Schulleistung stellte sich heraus.

Ein aus zehn Fragen bestehender Fragebogen zum Zweck der Auffindung von noch anderen Fällen der Linkhändigkeit wurde gegen die Testbatterie verwandt, und es erwies sich, dass er über 90% wirksam zur Auffindung derjenigen ist, welche linkhändig waren nach einem zusammengesetzten Wert, bei dem alle vier Tests ein gleiches Gewicht bekamen.

Durost

$7.00 per volume
Single numbers $2.00
This issue, $3.00

MONTHLY
Two volumes per year

Nov.-Dec., 1934
Volume XVI, No. 5-6

GENETIC PSYCHOLOGY MONOGRAPHS

Child Behavior, Animal Behavior, and Comparative Psychology

EDITED BY
CARL MURCHISON

Worcester, Massachusetts

Entered as second-class matter December 1, 1925, at the post-office at Worcester, Mass., under Act of March 3, 1879

$7.00 per volume
Single numbers $2.00
This issue, $3.00

MONTHLY
Two volumes per year

Nov.-Dec., 1934
Volume XVI, No. 5-6

GENETIC PSYCHOLOGY MONOGRAPHS

Child Behavior, Animal Behavior, and Comparative Psychology

AN EXPERIMENTAL STUDY IN THE PRENATAL GUINEA-PIG OF THE ORIGIN AND DEVELOPMENT OF REFLEXES AND PATTERNS OF BEHAVIOR IN RELATION TO THE STIMULATION OF SPECIFIC RECEPTOR AREAS DURING THE PERIOD OF ACTIVE FETAL LIFE

From the Psychological Laboratories of Brown University

by

LEONARD CARMICHAEL

Worcester, Massachusetts

Entered as second-class matter December 1, 1925, at the post-office at Worcester, Mass., under Act of March 3, 1879.

CONTENTS

I

GENERAL PROBLEM AND EXPERIMENTAL PROCEDURES

A. INTRODUCTION

In this monograph a report is given of a series of experiments which have been carried out by the writer during the last two years in an effort to gain a further understanding of the prenatal development of reflexes and of the origin and early growth of organized systems or patterns of behavior in a typical mammal. Specifically, the study deals with the responses released by the systematic stimulation of a large number of definitely located receptor areas, or "reflexogenous zones," in a series of fetal guinea-pigs each of known gestation age. The terms "reflex" and "behavior patterns" are defined in the discussion at the end of the paper. At this point it may be said that these terms are employed for convenience, and their use is not intended to imply any particular theory of behavior.

The conditions of the experiment are first described and then as complete a picture as possible is given of the complexity and yet the regularity of the behavior released by controlled stimulation of specific areas at various stages in these fetuses. At the end of this presentation the results of the study are summarized in a series of conclusions which involve certain positive statements and an evaluation of a number of current theories of the development of behavior. These conclusions are so presented as to show the bearing of the present investigation upon certain insistent problems of psychology.

B. The Experimental Animal

The guinea-pig was chosen for this study for a number of reasons. It is a readily available laboratory animal. Its fertility is such as to make it satisfactory for a study of fetal phenomena. Many important aspects of the morphology and physiology of the organism are known as a result of intensive study by biologists. Its long period of gestation seemed likely, on the basis of previous work, to present an unusually comprehensive series of fetal behavioral phenomena. And, finally, an excellent study of the development of behavior, from the psychological point of view, during the latter part of the gestation period, after air breathing could be established in the fetus, has been previously reported by Avery (6).

The work of Schulz in 1829, Bischoff in 1852, Reichert in 1861, Hensen in 1876, and Rubaschkin in 1905, on the special phenomena of reproduction in the guinea-pig, has been summarized by Stockard and Papanicolaou (86). Among the more recent writers who have contributed knowledge to this field and especially to the knowledge of the nature of the reproductive cycle in the guinea-pig may be mentioned Loeb (56, 57, 58), Read (78), Stockard and Papanicolaou (86, 87, 88), Ishii (49), Draper (28), Avery (5), Ibsen (45), Harman and Prickett (39), Young, Myers, and Dempsey (106), Harman and Dobrovolny (38), and Dempsey, Myers, Young and Jennison (25).

The animals used in the experiment were kept on a full diet and at relatively constant temperature condi-

tions. Characteristically, the mean length of the oestrous cycle is set at 16 days and 7 hours by Young, Myers, and Dempsey (106) in the adult female guinea-pig, the total range being between 13 and 21 days. There is variability also in the length of time that the "heat period" lasts. Its mean length is set at 8.01 hours by these same investigators. In the present experiment the colony of animals was examined frequently for behavioral and vaginal indications of sexual activity. When a female was found to be in heat it was placed with a sexually active male. The exact time of copulation was recorded. Every litter used in the present study was dated in this manner.[1]

In the guinea-pig ovulation almost certainly occurs within 8 hours after the end of oestrous. With safety one may assume that conception follows within 12 hours after copulation. Copulation age, with an error of less than 12 hours, thus gives the true age of the embryos of fetuses resulting from the known copulation. In this study, therefore, *copulation age* itself will be used in describing the age of fetal litters.

The period of gestation in the guinea-pig is generally taken as 68 days, although Draper and others have presented evidence that it may vary from 63 to 70 days. The wide range of time is possibly due to the confusion of late abortions with normal birth.

---

[1]The author wishes to express gratitude to Dr. William C. Young and to his collaborators, Messrs. H. I. Myers and E. W. Dempsey, who were kind enough to direct the mating of a large number of the animals used in the study. This work was done as part of a general investigation of the sex-cycle of the guinea-pig which is being carried out by Dr. Young and his collaborators.

In the present experiment fetal litters of known copulation age from 87 adult females were prepared for study. Because of accidents and the desire to use only absolutely normal animals, however, the data of this paper are based exclusively upon 60 litters which provided a total of 178 fetuses. In Table I the copulation age, an arbitrarily assigned identifying number, and the number of fetuses in each litter are presented.

TABLE I

SHOWING POST-COPULATION DAYS, IDENTIFYING NUMBER, AND NUMBER OF FETUSES OF ANIMALS ACTUALLY USED IN THE PRESENT STUDY

Days	No.	Fetuses	Days	No.	Fetuses
25	1	3	43	31	3
27	2	3	44	32	3
27	3	4	45	33	3
27	4	3	46	34	3
28	5	2	47	35	2
29	6	4	47	36	4
29	7	1	48	37	2
30	8	3	49	38	4
30	9	3	49	39	4
31	10	4	50	40	4
31	11	2	51	41	4
32	12	3	52	42	2
32	13	3	53	43	4
32	14	3	54	44	2
33	15	3	54	45	4
33	16	2	55	46	2
34	17	4	56	47	3
35	18	3	57	48	3
35	19	3	58	49	3
35	20	4	59	50	2
36	21	3	60	51	3
36	22	3	61	52	3
37	23	3	61	53	2
38	24	4	62	54	3
39	25	2	63	55	3
39	26	4	64	56	3
39	27	2	64	57	2
40	28	3	65	58	3
41	29	4	66	59	2
42	30	3	67	60	2

The difficulty in preparing this material leads the writer to confirm emphatically Draper's statement that the fecundity of the guinea-pig has been exaggerated.

C. Operative Preparation of the Pregnant Adult Animal

Preyer's study had indicated that the fetus of the guinea-pig was immobile until about the 28th day (76), as had that of Yanase (104). In the present study, therefore, it was planned to begin at a period just before this and then to study one or more litters dated on successive days up to the normal birth time, which was set at 68 days.

The experimental procedure consisted in administering ether for less than two minutes to the pregnant female whose fetuses were to be studied. As soon as the animal was completely anesthetized an incision of approximately 4 cm. was made in the mid-line of the dorsum of the neck parallel to the long axis of the body. The subcutaneous tissues and muscles of the neck were then quickly dissected away, exposing the cervical vertebrae of the spinal column. A specially constructed knife was then inserted between the 4th and 5th cervical vertebrae and the spinal cord completely transected. In most instances this whole operation required less than three minutes. Immediately after the operation, while the animal was still anesthetized, a long abdominal skin incision was made. The animal was then allowed a period of approximately 2 hours to recover from the effects of the anesthetic. As a re-

sult of the transection of the cervical cord just described, the maternal animal was, save for spinal reflexes, immobile below the region of the section. As judged by behavior, the whole region below the cut was insensitive. This operation was peformed because the decerebration of the adult female, as a preparation for the study of fetal behavior such as that employed, for example, by Coronios on the cat (22), proved difficult in the guinea-pig. The cord transection led to very little loss of blood and seemed to be a satisfactory operation for the purposes of the present study. It rendered the animal immobile and insensitive at the region of the abdominal operation, but did not interfere with the function of either lungs or heart and thus fetal circulation was maintained unimpaired.

After the period allowed for the passing of the effect of the anesthetic had elapsed, the prepared adult female was tied to a wire frame and lowered to the shoulder level into a specially constructed bath apparatus filled with physiological salt solution. The salt solution of the bath was maintained at a constant temperature of $37.5 \pm .5°$ C. by immersing the tray in which the solution was contained in a larger tank filled with tap water, which was kept at the desired temperature by thermostatically controlled electric heaters.

As soon as the adult animal, supported ventral side up at an angle of 30° on the frame, had been placed in the bath so that its body below the shoulder girdle was immersed in the warm salt solution, a deeper incision was made, under water, in the abdomen, and then by

a suitable further operation the fetuses were one by one released from the uterine membranes. Great care was taken to carry out this operation without loss of blood. Usually so little blood was actually lost that the water was not discolored. In this operation care was exercised to maintain the placenta, the umbilical cord, and the whole fetus itself in such a position that tensions did not arise which might interfere with normal fetal circulation.[2]

D. Experimental Procedures With the Fetal Organisms

Before opening the amnion a motion-picture camera fitted with special lenses was focused at approximately 20 inches above the bath apparatus, so that, as desired, a record could be secured of the first movements of the fetus made after opening the amnion and of significant subsequent sequences of behavior.[3]

A rubber-covered adjustable lead supporting-stand was arranged to provide a sort of hammock upon which the fetus to be studied could be supported in any desired position.

---

[2]The author wishes to express his gratitude to Dr. James D. Coronios who cooperated in each experiment carried out. The results of Dr. Coronios' observations on the same fetal material considered here will be published in a separate paper dealing with the motor development of the fetal guinea-pig in relation to age and external anatomical measurements. Thus, two (and sometimes three) trained observers watched every fetal movement which was recorded. They also checked the accuracy of each dictated protocol.

[3]The writer wishes to express his gratitude to the National Research Council for a grant in aid of research which made the purchase of this film possible.

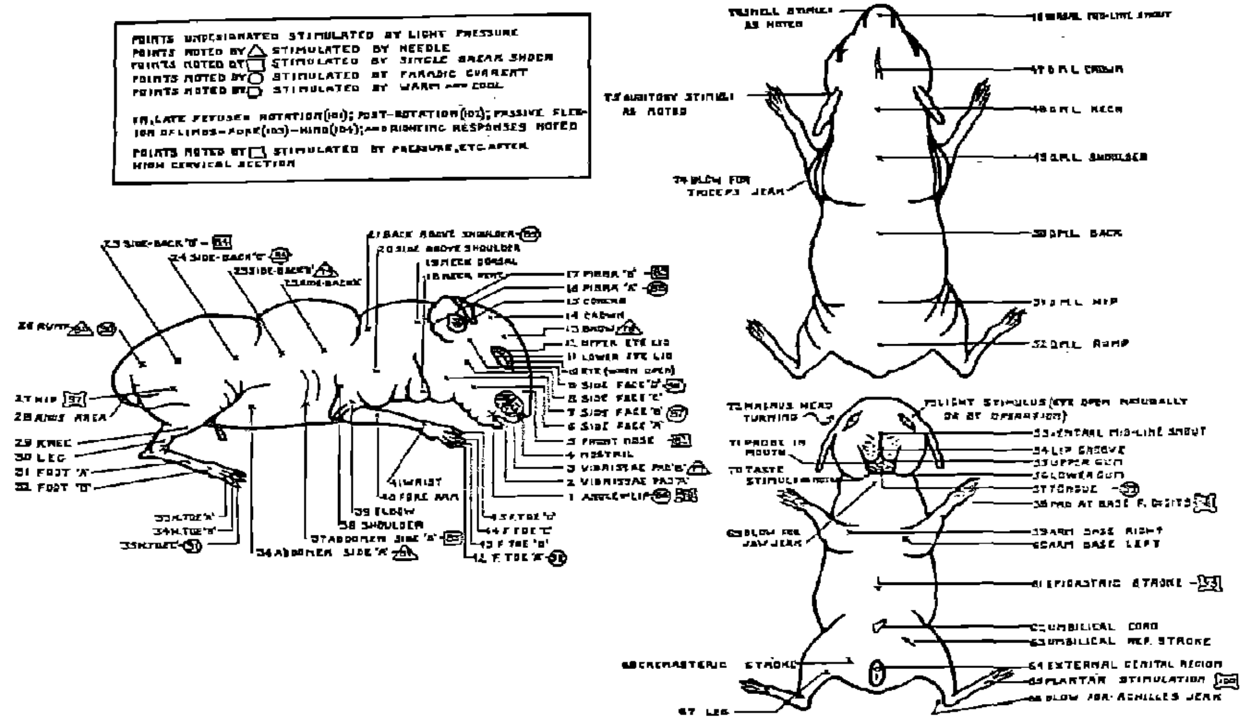


FIGURE 1

CHART OF RECEPTOR AREAS AS STIMULATED IN THE PRESENT STUDY

The points are indicated on a diagram of a late fetus. On younger fetuses the same relative areas were stimulated, save that, as explained in the text, an abbreviated procedure was necessarily sometimes used.

When the fetus had been thus prepared it was systematically studied by stimulating it at as many of the 104 points indicated on the Chart of Receptor Areas, given as Figure 1, as seemed expedient. To accompany this chart, which was used as a guide by the experimenter, a series of mimeographed eight-page booklets were made. One such booklet was used for each fetus studied. The description of behavior was written out in shorthand by the author's wife. Immediately after the experiment the shorthand notes were transcribed in permanent record booklets. During the course of each experimental period long sequences of behavior were recorded by the motion-picture camera. These pictures, of course, had to be taken of the fetus while it was under the water of the bath, and in order to make such photography clear two 500-watt spotlights were arranged so that they could be focused on the fetus while the camera was being used. The use of these recording films has made possible the observation of many aspects of behavior which would otherwise have remained unobserved. The advantages of this method of recording in the study of fetal and neonatal behavior have been pointed out elsewhere (91, 37).

At different times, as will be indicated in the study below, the behavior resulting from the following *modes of stimulation* was noted: (1) Non-experimentally controlled stimulation. This sort of stimulation is the result of general external and especially internal energy changes in the conditions affecting the organism or certain cellular systems of the organism and leads

to or facilitates behavior that is generally called "spontaneous movement." Under this heading the specific gravity and the normal temperature of the medium (such as air or water) in which the fetus is living should be included, as should also incidental stimuli such as those resulting from contact with the supporting frame, vibration of the room, and the like. For an evaluation of the concept of "spontaneous activity" see Skinner (84). (2) Light punctiform touch by a pliable hair fastened at right angles to the end of a wooden probe in the manner of a von Frey esthesiometer. (3) Light stroking by the esthesiometer, thus giving a form of areal stimulation. (4) Deeper pressure by a strong bristle or blunt probe. (5) Prick by a sharp needle also fastened at right angles to a wooden handle. (6) Light pinching of skin and superficial muscles by forceps. (7) Probes of various sizes to be inserted in the mouth. (8) Movement of an appendage at a joint by holding the distal end of the member and manipulating it. (9) Electrical stimulation by single-break shocks and by an interrupted current from an inductorium. (10) Drops of water of temperature the same as, above, and below that of the bath. (11) Loud sound produced by a whistle. (12) Bright light produced by 500-watt bulb in a spot-light. (13) Bends and twists of the head and neck in the study of "Magnus reflexes." (14) In late fetuses, removed from placental circulation and breathing air, rotation on a turntable. (15) Also in late fetuses taste and smell stimuli.

As noted in the report of results, the above modes of

stimulation were in many instances used at different intensities.

The responses given to the various forms of stimuli just outlined are not by any means to be considered as simple direct functions of the stimulus used. Indeed, a series of factors which may be called the *variable conditions of the organism* influence the results secured from stimulating active fetuses. Beside the mode and intensity of the stimulating energies described above, the response given by the fetus must be considered as a complexly determined act dependent as well upon the following factors: (1) The gestation age of the fetus and all of the correlated anatomical and physiological characteristics of the variable genetic stock and age in question, including, as most important, the structural development of the receptor-neuro-muscular mechanism. (2) The size of the area stimulated. (3) The number of stimulations given in a specific time (that is, allowance for the phenomena of "summation" and "fatigue"). (4) The transient physiological condition of the fetus at the time of stimulation (that is, dependence upon blood supply, immediately antecedent activity, and the like). (5) The general body posture of the fetus at the time of stimulation. (6) Unknown changes in the visceral or other internal conditions.

A consideration of the *variable conditions of the organism* just presented makes it clear that any simple correlation between a "single stimulus" and a "unitary response" in a fetus of a particular age cannot be hoped for. However, in all of the experiments of the present study every effort was made to maintain conditions

as constant as possible. It must be clearly stated, however, that, because of the different *modes of stimulation* and the *variable conditions of the organism* outlined above, absolute consistency on the same age cannot be expected. In spite of these variable factors, however, the writer has been surprised in the course of this study at the uniformity of response that appears in the whole series of fetuses.

The Chart of Receptor Areas shown in Figure 1 was developed after some preliminary experimentation. On the chart are indicated the zones which were stimulated in a routine manner in each fetus studied. *The novel results of this study are in no small measure due, it seems, to the fact that previously unrecognized behavioral sequences could be observed in this investigation as a result of the comparability of stimulus procedures from fetal stage to fetal stage made possible by the consistent use of this chart.* It is believed that this form of relatively complete, constant, and therefore comparable stimulation from fetal stage to fetal stage has never before been applied as systematically and extensively as it has in the present investigation.

II

EXPERIMENTAL RESULTS

The limitations of space prevent the publication of the full series of protocols taken during the course of this rather complete experimental exploration of stimulus-released behavior at each developmental level. *The full transcription of the protocols occupies more* than four hundred typewritten pages, and yet in this investigation, from one point of view, these complete protocols alone can give the true picture of fetal activity as it has been observed in the many fetuses of many ages used in this study. In the present paper a few complete protocols are presented which give the recorded responses in each of the series of fetuses studied resulting from the stimulation of a numbered receptor area. In most instances, however, only brief summaries of the protocols can be offered.

It should be emphasized that the course of experimentation was to stimulate *each* spot that could effectively be stimulated in *each* fetus studied. The complete record of response of each fetus was thus taken down on a record booklet, as described above. At no time did the observer read any previous protocol just before an experiment. The animals were also taken in random order, so far as fetal age was concerned. As much as 18 months sometimes intervened between, for example, the record of response to touching the lower eyelid in a forty- and in a forty-one-day fetus. Obviously suggestion and "knowing-what-to-look-for" were thus practically ruled out. *The consistency, or*

lack of consistency, of results given in any protocol as printed here is the result of reorganizing all the protocols and arranging the written observations in chronological order after the complete experiment was over. The consistency of results given in certain series of protocols was a complete surprise to the writer and was discovered when the study was being prepared for publication. The constancy, that is, became apparent only when the protocols which had, of course, been taken completely on each fetus were broken up and arranged chronologically according to each receptor area studied.

In a previous paper the present writer has discussed the problem of the stage at which observations which are to be taken as significant for an understanding of behavior should begin (15). The succession of generations is endless. Always in dealing with an animal some antecedent life events may be discovered which are significant in understanding the organism as it is at the time of observation. This is true whether it be the condition of the germ cells before fertilization or the make-up of the response mechanism before the first reaction to internal or external stimulation has occurred. For a scientific and logical orientation in this field see Gesell and Thompson (37).

In the present paper, rhythmic muscle contraction of the sort characteristic of the beating heart of the embryo will not be considered as the beginning of true behavior, but rather the first externally observable response of skeletal muscle will be taken as marking the onset of such activity.

In the present study the first response which could be surely identified occurred in an organism of 28 post-copulation days. In this stage, and indeed in the five or six days immediately following this first observable external movement, it was not possible to use the systematic *Chart of Receptor Areas* which has been described above. The first transitory behavior phenomena may therefore best be summarized in the following paragraphs under heading "A", while under heading "B" will be given the much more extensive series of protocols taken on animals which were systematically investigated.

A. Reaction of Early Motile Stages (27-31 days inclusive)

27 Days. Three separate adult animals were used in studying this age (Animals Nos. 2, 3, and 4). In no one of the nine normal fetuses studied was there any externally observable movement. Regular rhythmic heart beat could be observed in each fetus through the thin and almost transparent chest wall. Faradic stimulation led to direct muscle contraction.

28 Days. The animal (No. 5) studied at this age had two fetuses. No movement was observed in the first fetus, although muscle contraction was elicited in it by faradic stimulation resulting from applying electrodes directly to the muscles. In the case of the second fetus, while still inside the wholly transparent amnionic sac, "spontaneous movements," in the sense of the term as defined above, could be observed. These movements consisted in a slow lateral bending of neck and possibly thoracic trunk on one side so that the head was moved laterally. Weak flexion of the fore leg was observed at the same time. It seemed that this fore-leg movement once occurred without any accompanying trunk or neck movement. When this organism was allowed to come out of its sac into the warm bath of salt solution no further movement was observed, nor could movement be elicited by cutaneous stimulation.

29 Days. Two adult animals were studied at this stage (Nos. 6 and 7). The first had four fetuses. The first of these fetuses demonstrated spontaneous movements before removal from the sac. The most notable of these movements was a synchronous bending of the neck so as to bow the head at the same time that a definite flexion of the most readily observable fore leg occurred. After removal from the sac no further movement, save heart beat, could be observed. There was no response to exteroceptive stimulation. In the second fetus movement was seen as the fetus was shelled out of the amnion. This fetus gave marked direct muscle contraction to faradic stimulation. In the third fetus, while still in the sac, movement of the head and flexion of both fore and hind limbs were observed. When removed from the sac, however, no further movement could be observed. In the fourth fetus no movements were observed, save those resulting from direct electrical muscle stimulation.

The second adult animal operated upon at this age had but one fetus. No movements save those brought about in muscles by direct faradic stimuli were observed. The heart beat of this fetus was regular, however.

30 Days. Two adult animals were operated on at this age (Nos. 8 and 9). One of the animals had three fetuses. The first fetus of this animal, observed while still within the amnionic sac, showed spontaneous definite flexion movements of the fore limbs. These movements seemed to involve both head and trunk responses and probably also slight flexion of the hind limbs. These movements continued after the fetus had been removed from the sac, but were of smaller amplitude. No reactions could be clearly shown to result from stimulation applied by the experimenter. In the second fetus of this litter, also observed in the sac, flexion of the fore and possibly also of the hind limbs was observed, accompanied by lateral head movement. No definite trunk movement could be observed. In the case of the fore-limb movement it seemed that there was slight articulation at the elbow. The third fetus of this litter, after being shelled out into the bath, showed a unilateral flexion of the trunk (a "C" movement) so that both head and rump were brought nearer the surface of the water. Fore-limb flexion was observed at the same time with flexion at elbow joint. The hind-leg movement which was observed was possibly related to the bending of the body. In the case of each of these fetuses the time of the observation was

limited. Movements persisted only from a few seconds to two minutes after observation began.

The second adult operated upon at this period had three fetuses. In the first of these, while still in the amnionic sac, the left fore leg was observed to be brought up to the side of the face. Heart beat was also observed, but no further movement. No movement was seen in the second fetus of this litter. While the third fetus was still inside the sac, a sharp spontaneous movement of the left fore limb, adducting it toward the belly, occurred. A general lateral trunk flexure also occurred in this fetus. In no one of these three fetuses could movement by stimulation be elicited after it had been removed from its sac.

31 Days. Two adult animals were studied at this stage (Nos. 10 and 11), but in the case of one animal no movement or heart beat could be observed in any of the four fetuses. In the second adult animal no movement or heart beat was observed in the first fetus exposed, but rather elaborate "spontaneous" behavior was recorded for the second fetus. In this fetus, after removal from all sacs, definite lateral trunk flexion was observed. These movements were accompanied by fore- and hind-leg flexion and by weak neck flexion leading to bowing of head. At another time a number of strong "C" movements brought both head and rump toward surface of water. As part of this "C" movement both hind legs moved with what seemed to be articulation not only at the hip but at the knee. At one time the "C" movement seemed to lead to an extension of the spine with a consequent postural stiffening of the whole trunk. What seemed to be slight contraction of the muscles of the abdominal wall was also observed. In one instance a slight circumduction of the fore limb followed by a distal movement of the limb with articulation at the carpus took place. Again and again hind-leg flexion followed by flexion of the contralateral fore leg was observed. The fore-limb movements, however, were smoother and more definite than were the observed hind-leg movements.

This 31-day fetus was the youngest fetus of those observed in this study to give a definite response to experimentally applied external stimuli. When touched on the face no definite response could be elicited, but when the esthesiometer was applied to the region which would later form the fully developed concha of the ear definite fore-leg movements were elicited which

were accompanied by a flexion of the neck which led to a bowing of the head. The bowing occurred synchronously with the slight raising of the fore limb. A motion-picture record of this movement was fortunately secured.

B. Reaction of Fetuses Showing Definite Responses to Experimentally Presented Stimuli

In the paragraphs above we have seen that, as the result of conditions independent of any stimuli applied by the experimenter, quite elaborate responses occur. In the case of one fetus of 31 gestation days just considered, it proved possible to bring about response to tactual stimulation.

On the following pages are presented a series of observations concerning the onset and change in responses as such responses appear following the systematic exploration of the exteroceptive field of a series of fetuses as shown in Figure 1.

In each case the responses given are the result of lightly applied pressure as described above, when the contrary is not specifically noted. In the case of the complete protocols a series of numbers in brackets at the beginning of each observation signify (1) the age in days of the fetus determined from the recorded time of copulation, as D32, (2) an arbitrarily assigned identifying number referring to the adult female in which the fetus developed, as No. 13, and (3) in certain instances, a number preceded by F, which indicates a particular fetus of the litter, as F2. (Where no F is given it is assumed that Fetus 1 is being considered.)

The letters N. R. signify that no response followed stimulation.

In reporting the results, the point of stimulation as defined on the chart will be presented and then the reaction to this point will be given at successive gestation ages from 32 days after copulation to the end of the normal prenatal developmental period at 68 days. It must, of course, be remembered that in younger stages it was not possible to stimulate at all the points indicated upon the chart, because anatomical development had not progressed to a point at which such fine discrimination of area was in all cases possible. It must also be kept in view that the responsive life of the early motile fetus is not as great as it becomes in the middle and later part of the gestation period. Thus, it was not possible to explore the receptive areas of the 35-day fetus for more than a few minutes before definite signs of a decrement in response were noted. On the other hand, in later fetuses it was possible to study one fetus for several hours and still have no apparent loss in the speed or adequacy of response. It must also be apparent in these reports that the regions which were selected before the experiment began as points of stimulation are, from the results of the experiment, shown to be in some cases needlessly overlapping with other points, and in other instances the points selected were rather non-significant. Complete mapping of the reflexogenous zones which are indicated here would necessarily require a psychophysical technique similar to that employed in mapping the blind spot of the eye. Such a procedure would entail the use of dozens or

even hundreds of animals at each gestation age. Once again, therefore, it may be said that the results given below should be taken as preliminary rather than as the result of a final survey of the growth of the reflexogenous zones and stimulus-released behavior in the fetal guinea-pig.

We now turn to consider, area by area, the responses released by systematic stimulation at different post-copulation ages.

1. *Angle of Lip.* (D32, No. 13). A slight lateral flexion of neck occurred at once, thus bringing head toward the stimulus. At the same time a lateral flexion of the rump took place. The fore leg was then brought up toward but not touching the mouth. (D32, No. 14, F2) At first both fore paws were raised toward the head and then synchronously flexed. (D32, No. 14, F3) The fore paw was raised so as actually to touch the point stimulated. After a few seconds the stimulus was repeated and exactly the same precise localizing movement occurred. (D33, No. 16) A slight extension of the back occurred, which was quickly followed by a flexion of the whole trunk and a flexion of both hind legs. (D34, No. 17) The head was pulled up and rotated sharply so as to remove point of stimulation from stimulus. Both fore legs were then brought up toward snout; the homolateral leg brushed the spot stimulated during this movement. On re-stimulation the same sequence occurred. (D35, No. 20) The head was pulled away to side. A twitch of all four limbs occurred, which was followed by a definite but slow raising of the homolateral fore leg up to side of face. In this case the paw did not touch the face. When the paw was over the face, however, it executed several "brushing movements." (D35, No. 19) After stimulation an opening of the mouth occurred at once. This was followed by responses of the body muscles which included fore- and hind-limb movements. (D36, No. 21) A shake of the head with extension of trunk occurred at the same time that a rather rapid raising of homolateral fore limb took place. This posture of the fore limb was maintained for several seconds. (D36, No. 22) The homolateral leg was brought up to mouth.

During this time the mouth opened and closed several times. Eventually several digits of the paw were placed in the mouth. (D37, No. 23) Stimulation led to movement of the homolateral fore leg toward the angle of the lip. (D37, No. 23, F2) A very light touch of lip led to slight pushing out of the tongue (D38, No. 24) The homolateral fore limb was extended. On second stimulation this was repeated, but in neither case did the paw come in contact with the face. (D39, No. 27) On light pressure stimulation the tongue was pulled into the mouth. The muscles of the vibrissal area contracted so as to corrugate the skin above them. Fore-leg movements were executed which involved brushing of the snout. (D39, No. 25) The vibrissae pad was slightly corrugated and the heterolateral hind leg twitched and was then brought up toward but not touching the point of stimulation. (D40, No. 28) A twitch of the vibrissae pad to a single touch was noted. To repeated stimuli the twitch of the vibrissae pad occurred again and both fore legs were brought up and the homolateral limb was brought over so that the spot was brushed. (D41, No. 29) The tongue was thrust forward. (D42, No. 30) Very light stimulation led to protrusion of the tongue. (D43, No. 31) The homolateral fore limb was brought up to the spot, which was then brushed by the paw. (D44, No. 32) To a very light touch, a slight opening of the mouth with a corrugation of the vibrissae pad occurred. To a stronger stimulus a slight upward snap of the head took place, resulting from neck extension, and both fore legs were brought up toward the face. Tongue movements also occurred. (D45, No. 33) Specific and definitely localized movements of the upper lip occurred on the side toward the stimulus. At the same time the homolateral fore leg was flexed at the elbow and the limb was brought up to the snout with a maintained flexion also at the wrist. The digits were slightly flexed and maintained in a "cupped" posture. The homolateral hind leg was flexed and adducted toward the belly. A slight twitch of the head also occurred. (D45, No. 33, F2) Elevation of both the upper lips in a "snarl" took place accompanied by a quick sharp extension of the spine, thus stretching the whole body. This occurred at the same time as an upward jerk of the head. An extension of both hind and fore legs also took place. The fore legs were then brought up to but did not touch the snout. *(D46, No. 34) To stimulation a quick twitch*

of the vibrissae pad occurred. Movements of the fore legs in the general direction of spot stimulated then took place, but the reaction was not specifically localized. (D47, No. 35) On stimulation both fore legs were rapidly extended and elevated so that they touched the lower part of the chin. When a small brush was substituted for the single-hair stimulus the mouth opened and the upper lip was raised with a precise local response. (D48, No. 37) To stimulus a twitching of the vibrissae pad could be observed, together with an extension of the homolateral fore limb, which did not rise far enough to touch the point stimulated. (D52, No. 42) After stimulation, movement of all four limbs was noted, with possibly a greater flexion of the homolateral fore limb than of any of the others. (D53, No. 43) Stimulation led merely to a short twitch of the vibrissae pad. (D54, No. 44) At first to stimulation at the exact point indicated on the chart no response occurred, but by moving the hair a few mm. to the outer edge of the vibrissae pad a single touch led to a twitch of the paw on that side. When the angle of the lip was again stimulated, the mouth opened with a contraction of the muscles around the angle of the lip and a protrusion and retraction of tongue occurred several times. (D54, No. 45) To stimulation a contraction of the upper lip occurred, including a corrugation of the vibrissae pad. (D55, No. 46) After stimulation a local contraction of lip took place, which was followed by a slight flexion of neck, leading to head movement. This last response was accompanied by the raising of, first, the homolateral and, then, of the heterolateral fore limb. Eye wink also occurred. (D56, No. 47, F1) No response to stimulation in this area, no matter how repeated; fetus was fully active in other respects. (D56, No. 47, F2) Contraction of pinna on the same side followed stimulation. The heterolateral fore leg was also brought up slightly and the heterolateral hind leg brought up and fluttered rhythmically as in a "scratch response." (D57, No. 48) Quick precise opening of the mouth occurred as a result of stimulation. This was synchronized with a toss back of the head in such a manner that the movement became exactly that of a dog catching a thrown ball in its mouth. (D58, No. 49) At first no response followed stimulation. Continued stimulation, however, led to lip movements and a quick bite, followed by continued lip movements. Then there was a general trunk extension, changing the orientation of the whole body in space. (D59, No. 50) No re-

sponse even to continued stimulation. (D60, No. 51) A jerk back of the head was the first response to stimulation. This was followed by a rhythmic scratch reflex of the homolateral hind leg and then by a scratch response in the heterolateral hind leg. Flutter of both fore limbs now began and continued during several rhythmic extension-flexion responses. (D61, No. 52) The homolateral hind leg was extended and elevated to the shoulder in a series of rhythmic scratch movements. At the same time the head slightly rotated and extended in a manner to make the stimulus more intense. This was the opposite of a so-called "avoidance" action. After a few seconds' rest, this stimulus was given again with identical results. (D61, No. 53) N. R. (D62, No. 54) To light touch stimulation a sharp pulling away of the head occurred, followed by an extensor thrust of the fore paws from the body. (D63, No. 55) As a result of the application of the stimulus a slight localized movement of the upper lip took place. No other response occurred. (D64, No. 56) There was a slight movement of the head. This was followed by an extension and flexion of the heterolateral fore limb, and then by a similar movement in the heterolateral hind limb, and then by similar responses in the homolateral fore limb and homolateral hind limb. On stimulating again, this series of leg movements was repeated as a trot response. All of these movements were of very slight magnitude. (D64, No. 57) N. R. (D65, No. 58) As a result of a touch stimulation a twitch of the trunk muscles took place which also seemed to involve the fore and hind limbs. (D67, No. 60) To stimulation a slight twitch of the head took place. The homolateral fore leg was then thrust up to the snout. The homolateral hind leg was extended and adducted to side just below the shoulder and, while in this maintained posture, executed several rhythmic scratches.

2. *Vibrissae Pad "A"*. (D32, No. 13) A slight extension of the neck occurred, thus tossing the head back. At the same time a twitch of slight amplitude occurred in both fore legs. This exact sequence of response was given again to a second stimulus. (D32, No. 12) Stimulation led to a raising and adduction of the fore leg so that the paw swept over the point stimulated with perfect precision, as if to brush away the stimulus. (D32, No. 14) A double lateral flexion of the spine followed stimulation. It began at the cervical region and proceeded caudad to form the typical "sigmoid

movement" of Coghill. (D33, No. 16) Definite homolateral fore-limb raising and adduction to the face resulted from the stimulus. This movement brought the paw toward but not touching the point of stimulation. At the same time the neck was slightly extended, thus lifting the head. (D34, No. 17) The head was sharply rotated, thereby removing the point of stimulation from the posture which it had held when the stimulus occurred. The homolateral *leg was first raised and then flexed and adducted to the belly. (D35,* No. 20) To a single touch, a slight twitch of the head and of the homolateral leg took place. To a slight stroking stimulus, the head was flexed toward the side of stimulation. At the same time the fore leg was extended and adducted to the side. (D35, No. 19, F3) Stimulating the point indicated led to violent movements of the head such that the crown of the head, as a result of lateral flexion of the neck, came toward the stimulus, while the snout moved so as to avoid the stimulus. Repeated stimuli gave this same sequence of behavior several times. On one occasion the pattern of response was complicated by the fact that the homolateral fore paw was brought up strongly to the snout. Stimulation led to violent head extension and flexion on the part of the animal. These movements were accompanied by rhythmic paw flexions and by movements of the lower jaw so as to open and close the mouth. Only after five or six such movements did the animal bring its paw up as if to brush away the stimulus. The movements antecedent to the brushing movement involved reactions in many other directions. Following this one stimulation for a period of half a minute after the first brushing movement the fetus continued to "scrub" with its paw at the point of stimulation. (D36, No. 22, F2) In response to stimulation a slight jerk of the head took place followed by a twitch of the fore limbs. (D37, No. 23) No response to a single touch, but continued stimulation led to responses involving trunk and limb muscles. On stimulating by a single touch again, after an interval of a few seconds, the homolateral fore limb was slightly extended and moved toward point of stimulation, but not so that the paw touched the spot. (D38, No. 24) As a result of a single touch the neck was sharply flexed, thus leading to a bowing of the head. The homolateral fore limb was raised sharply with an accompanying twitch of the trunk. After a delay of a second, but with no further external stimulation, both fore legs were ex-

tended and adducted so that the vibrissae pads on both sides of the face were brushed, a definite contact being established between each paw and the face. (D39, No. 27) Corrugation of the skin of the vibrissae pad as a result of local contraction of the muscles just below it resulted from stimulation. This was followed by a slight extension of the homolateral fore leg. (D39, No. 26) Lightly touching the spot indicated elicited a sharply localized contraction of the musculature under the pad. Both fore legs were then simultaneously flexed and pulled back toward the hind legs. The hind legs were brought up and at the same time extended with articulation at the hip joint. (D40, No. 28) Stimulation elicited a twitch of the muscles below the vibrissae pad. (D41, No. 29) Corrugation of the pad as a result of contraction of the muscles under the pad occurred as a result of stimulation. At the same time a slight toss of the head resulted and the fore limbs were both slightly extended. (D42, No. 30) Stimulating the point noted led to a jerk of the head up toward the stimulus which had been applied. No contraction could be elicited when, instead of touching the skin, the short vibrissae themselves were bent by the aesthesiometer. (D43, No. 31) Twitching of the musculature of the pad occurred to stimulation. The homolateral leg was brought up and brushed past the spot. (D44, No. 32) To a light touch there occurred a slight contraction, that is, corrugation, of the vibrissae pad. To a stronger stimulus, an extension of the trunk and an extension and elevation of both fore legs occurred. (D45, No. 33) Stimulation was followed by a slow twitch of the muscles under the vibrissae pad and by a very slight twitch of the fore legs. (D45, No. 33, F2) A twitch of the vibrissae pad, opening and closing of the mouth, and short convulsive flexion, synchronous with a lateral twisting flexion of the trunk, took place as a result of stimulation. (D46, No. 34) Stimulation led to a twitch of the vibrissae pad and to a slight flexion of the neck muscles so as to move the head away from the stimulus. The homolateral fore paw was brought up to the side of the face, but did not touch the exact spot stimulated. Following this, the trunk executed a series of writhing movements. (D47, No. 35) Stimulation of the pad was followed by movement of the lower jaw. A second stimulation, however, led to no movement. A third stimulus was followed by a slight movement of both fore paws. After this the stimulus hair was used to touch the vibrissae

without coming near the skin at their base. This stimulation led to a rapid extension of the homolateral paw. On repetition of this stimulation, both fore limbs were rapidly extended and by a quick movement brought up to the nose. A third stimulation of the same sort led to very similar results, although there was a greater latency of response time. Following this stimulation the pad was again touched and the fore limbs were moved up. Thirty seconds after this stimulation of the pad of the vibrissae the fore paw was brought up to the side of the face. The carpus was then flexed, and this movement, together with the limb flexion, led to a brushing of the cheek about 5 mm. from the point of stimulation. (D47, No. 36) Stimulation led to local movement of the muscles at the base of the vibrissae and then to extensor thrusts of both fore limbs. (D48, No. 37) Contraction of the muscles below the vibrissae pad with very precise localization followed stimulation. (D49, No. 38) Stimulation led to extreme flexion of the neck, thus bowing the head almost to the chest. This was followed by an extension and circumduction of the heterolateral fore limb with marked flexion at the carpus so that the back of the paw brushed the exact point stimulated (that is, the vibrissae pad on the side opposite to the active paw was scratched by that paw). After this the homolateral limb was raised and the paw brought into contact with the point stimulated. During this activity the hind limbs carried out rhythmic "scratch reflex" movements. *(D50, No. 40) Local contraction of lip and vibrissae* pad followed stimulation. Slight fore-paw flexion and mouth movement were also elicited. (D51, No. 41) To a light touch stimulus the mouth opened and closed and the tongue protruded. The neck was extended so as to toss the head back and away from the stimulus. Both fore legs were brought up and the paw so flexed that the exact point stimulated was brushed. The muscles underlying the pad twitched locally as a result of the stimulus. (D52, No. 42) Sharply localized contraction of the muscles under the pad took place. The tip of the nose was pulled away from the stimulus by the contraction of some of the facial muscles. On repeated stimulation a quick flexion of the fore paw took place. When great care was exercised so that one or two of the vibrissae alone were touched, that is, slightly bent, a response of greater amplitude occurred than to actual stimulation of the pad. This response included local mus-

cle reactions and fore-limb extension. (D53, No. 43) N. R. (D54, No. 44) Stimulation led to contraction, that is, corrugation, of the musculature below the pad. (D54, No. 45) Stimulation led to corrugation of the musculature below the pad. (D55, No. 46) Local contraction of muscles underneath vibrissae pad occurred to stimulation. Then followed wrinkling of the snout, a shake of the head, and a quick extension and elevation of the paw so that it brushed the nose. (D56, No. 47) N. R. (D56, No. 47, F2) Stimulation led to sharp contraction of the eyelids, vibrissae pad, and of the snout, and a pulling away of the head as a result of dorsiflexion of the neck. Fore-leg paddling movements with the legs gradually more and more elevated until they came nearer the nose were then executed. (D57, No. 48) Just perceptible corrugation of skin at base of the vibrissae took place to stimulation. (D58, No. 49) Muscles of the pad twitched to continued stimulation, although at first there had been no response. (D59, No. 50) N. R. (D60, No. 51) No response occurred to repeated touch stimulus. A slight touch of the needle, however, led to a quick flexion and extension of both fore and hind homolateral limbs. Then, as a result of a virtual "contortion," the knee of the homolateral fore leg was brought up to actual contact with the point stimulated. (D61, No. 52) No response to light or heavy pressure stimulus. Slightest needle touch, however, led to a head shake, fore-leg extension, hind-leg flexion, and eventually adduction and elevation of the homolateral hind leg to the side. Mere touch of the vibrissae, however, after this response did lead to slight extension of the neck and extension and elevation of both fore legs. (D61, No. 53) N. R. (D62, No. 54) A slight lateral shake of the head took place, followed by a raising of both fore legs and a marked and maintained extensor thrust of the hind legs. (D63, No. 55) N. R. (D64, No. 56) No response to light pressure, but deep pressure led to flexion of the heterolateral fore limb, followed by a flexion of the homolateral hind limb. The same response was elicited by bending one or two vibrissae hairs. (D64, No. 57) N. R. (D65, No. 58) N. R. (D67, No. 60) To a single stimulus there occurred the sharp closing of the previously open eye. The mouth also closed. A localized pinna twitch then occurred, followed by a jerk of the head resulting from a quick extension of the neck muscles. Both fore legs were extended, ele-

vated, and adducted, so that they each brushed the homolateral side of the nose. The hind legs were adducted to the belly. At the same time the whole cervical part of the trunk was circumducted without moving the caudal trunk. A local corrugation of the skin below the vibrissae pad was also noted.

3. *Vibrissae Pad "B"* (Summary only). The responses elicited by stimulating this area were remarkably like those secured in the protocol just given. Corrugation of the pad appeared one day earlier (D38, No. 24). The protocols do not give any strong reason to suppose that the whole vibrissae pad may not be considered as a unified reflexogenous zone. Eye wink was, however, more frequently elicited and tongue movements somewhat less frequently elicited by stimulating area 3 than area 2.

4. *Nostril.* (D32, No. 13) Stimulation led to extension of the neck muscles and consequent upward toss of the head. The head was then brought down and the homolateral fore leg raised so that the palmar surface of the paw brushed the side of the face. A second touch led to a slight upward movement of the fore limb toward the spot but this time it did not touch it. To a more vigorous touch at the same spot a response occurred which included a sharp sudden lateral twitch of the fore leg, a lateral flexion of the trunk toward the stimulated side, and a flexion and adduction of the homolateral hind leg. A third stimulus led to a sharp lateral flexion of the neck, thus jerking the head away, while both fore legs were raised and adducted exactly to the place of stimulation. (D34, No. 17) To a light touch stimulus the neck muscles were extended, thus pulling the head upward. At the same time a slight backward kick of both hind legs took place. Both fore legs were extended, but the homolateral leg was not only extended but also elevated and adducted to the face, so that the point of stimulation was brushed by the fore paw. Repeating this stimulus led to a repetition of these responses, but the movements were less in amplitude. (D35, No. 20) The heterolateral fore limb was extended and elevated, but did not touch the point stimulated. A unilateral flexion of the trunk moved the head away from the posture in which it had been when stimulated. Both hind legs were flexed. (D35, No. 19) Both fore legs were raised so that both paws came in contact with the snout. (D36, No. 21) Stimulation was followed by neck extension and extension and elevation

of both fore paws. Brushing movements over the snout were then carried on by both fore limbs. Flexion of both hind limbs and body writhing also took place. After a period of quiescence without further stimulation both fore limbs were brought up in such a way that the tips of the digits were brought in contact with the inner surface of the lips. (D36, No. 22) Extension of the neck followed stimulation, thus jerking the head upward and away from the stimulus. (D37, No. 23) No response to slight stimulus, but continued stimulation led to a slight flexion of the neck and head. This was followed by continuous "scrubbing" movements with both paws partially flexed. This continued for some time and was in duration and thoroughness of "brushing" a novelty to the observers. (D38, No. 24) The homolateral limb was extended and adducted to the head slowly after stimulation. The limb was flexed at the wrist, and the digits were extended so that they came in direct precise contact with the exact spot that had been stimulated. (D40, No. 28) Stimulation led to the sharp lateral flexion of the neck, thus jerking the head away sharply in such a manner that the neck and body were almost at right angles. Extensor tonus in the fore limbs was maintained as they were brought up to the snout by a series of pushes. (D41, No. 29) The homolateral paw was brought up toward the nose after stimulation. The mouth then opened, the tongue was protruded and retracted rhythmically, and the head extended. At the same time the homolateral hind leg was flexed and adducted to the belly. (D42, No. 30, F2) Stimulation led to an extension and elevation of both paws toward but not touching the nostril region. When the limb was thus extended, the posture was maintained and nose touched by flexure at carpus. If right nostril was touched, flexure was of right wrist; if left was touched, flexure was of left wrist. (D42, No 30, F3) On stimulation both legs were extended and sharply adducted so that the paws almost met behind the head. The digits were then slightly flexed and the paws scraped down touching both sides of the face all the way to the snout. (D43, No. 31) Circumduction of the head occurred while both fore legs were extended and elevated and then flexed so that the snout was brushed. (D44, No. 32) No response followed light touch. Strong touch, however, led to an extension of the neck muscles and consequent backward toss of head which was followed by a twitch

of the fore legs. (D45, No. 33, F1) No response even to repeated stimulation. (D45, No. 33, F2) Neither response to touch nor at first to light touch of sharp needle was secured. A second touch of the needle, however, led to an extension of the fore limb slightly flexed at the elbow and at the carpus and with pronounced flexure at the digits. The limb in this posture was raised almost to the tip of the nose but not touching it. (D46, No. 34) Stimulation led to a twitch of the vibrissae pad, a wrinkling of the snout, an eye wink, violent flexure and rotation of the head so that the point stimulated was quickly removed from the locus in which it had been stimulated. Violent alternating paddling thrusts of both fore legs also took place, and after some seconds the trunk executed several elaborate writhing movements which seemed to involve both lateral and dorsal flexion and extension. During the course of this activity the hind legs were brought up to the neck. (D48, No. 37) Twitching of the nasalis muscles followed stimulation. The homolateral fore leg was then extended, elevated and adducted so that as it was gradually flexed the paw rubbed against the whole side of the face. (D49, No. 38) Local movement of the nasalis muscles which then seemed to spread so as to include the lip and tongue muscles followed stimulation. The tongue also executed local movements. The heterolateral limb was then brought up so as to brush approximately the exact point stimulated. The homolateral limb was then raised and then rubbed down the whole side of the face. During these responses the caudal region of the trunk was very active, and the trunk was so completely flexed that the hind limb virtually came in contact with the point stimulated. Following and during these responses there was marked activity of the general somatic muscles. (D50, No. 40) Stimulation led to mouth opening; then the homolateral paw was brought up toward the head and then the neck muscles were extended so as to pull the head backward. This extreme body tonus was maintained for some time. (D51, No. 41) The head turned sharply away after stimulation and both fore legs were brought up so that both paws brushed by the nostril. The trunk executed an "S" flexure, thus removing head from the locus of the stimulus. This response initiated violent rhythmic trot movements of both fore and hind legs which actually propelled the animal through the water as far as the umbilical cord would allow.

Mouth movements, tongue movement, and an eye wink also took place. On a second stimulation there was also much activity of trunk and limb muscles and this time both fore paws were brought up and held over the nostrils with flexed digits. (D52, No. 42) To a slight touch there occurred a toss of the head with accompanying trunk extension. On continued holding of hair in the nostril, the neck was flexed so that the head bent forward and all four limbs extended. The homolateral fore limb was partly flexed at the shoulder, elbow, wrist, and at the digits joints. Thus, the tip of one digit was brought directly to the point stimulated. (D53, No. 43) Digits of both fore paws flexed after stimulation. After second stimulation the muscles under the vibrissae pad and the nasalis muscle contracted. (D54, No. 44) A light touch led to a slight shudder of the head with accompanying twitch of both fore legs and a dilation of the nostrils. To somewhat stronger pressure the head was flexed and turned toward the stimulus, the mouth opened and closed, and the homolateral fore leg brought up so that it brushed the side of the face. (D54, No. 45) Stimulation led to a slight flutter of the fore legs and, when tried again, to a dorsiflexion of the head. (D55, No. 46) Stimulation led to a dorsiflexion of the neck muscles, thus quickly raising the head and pulling the nostril away from the stimulus. This was followed by a strong extension of the trunk and by quick "paddling movements" of the fore limbs. During this action there was a slight flutter of the hind limbs. (D56, No. 47, F1) Eye wink and a small, weak upward extension of homolateral leg followed stimulation. This same sequence was repeated on a second application of the stimulus. (D56, No. 47, F2) Stimulation led to a rotation of the head away from the stimulus. At the same time the heterolateral fore leg was brought up to and touched the side of the neck. (D57, No. 48) To the very lightest touch the eye closed and maintained the closed posture for several seconds. At the same time the pinna quickly twitched and there was a slight extension of both fore legs. (D58, No. 49) Stimulation followed by closing of the eye on the same side, flexion of neck muscles, and consequent head bowing, while the paw was brought up near the head. In this animal, when the eye was open, any slight touch stimulus on any part of the head seemed to lead to closing of the eye. (D59, No. 50) Corrugation of the vibrissae pad and nostril

and a twitch of all four legs followed stimulation. When the stimulus was given again, the heterolateral fore leg was raised toward the place stimulated. (D60, No. 51) After stimulation the muscles of the side of the face contracted and paddling movements of both fore legs were initiated. The homolateral fore leg was not only extended but elevated, so that the fore paw scratched the place that had been stimulated. (D61, No. 52) Head jerked away sharply to stimulation. Both fore legs were then thrust out and extended away from the chest. At the same time the hind legs were flexed, and the homolateral hind leg executed rhythmic scratching movements. (D61, No. 53) To stimulation the head was turned away. The homolateral fore leg was brought up to the side of the face and the homolateral hind leg was adducted so that with a twist of the trunk the whole fetus once pulled itself off the observing stand. (D62, No. 54) Rabbit-like movement of the nasalis muscles followed stimulation, as did a shake of the head and a slight abduction of fore paws. Eye-movements, pinna twitch, and lower-jaw movements also were recorded. (D63, No. 55) There was a localized contraction of the muscles lying beneath the nose. Both fore paws were then brought up toward the nostril. This took place again and included a closing of the previously partly open eye and a movement of the limb with paw and digits cupped. This response led to touching the spot that had been stimulated. (D64, No. 56) A twitch of the nose muscle followed by an extenson of the homolateral fore paw with flexed digits and a flexed wrist. The limb was brought up definitely so that scraping past the point just stimulated was accomplished. (D64, No. 57) There was a localized twitch of the nasalis muscles and a slight corrugation of the vibrissae pad following stimulation. (D65, No. 58) Light touch led to a pronounced flexion of the neck and consequently a pronounced dip of the head. This was followed by the extension of the homolateral fore limb which was brought up to the snout with the wrists and digits both flexed. The hind legs were flexed and adducted to the belly. (D67, No. 60) The head turned away from the stimulus after a touch, and a pinna twitch, eye wink, and a rapid thrust of both fore legs up against the snout also followed.

5. *Front Nose* (Summary only). Although the fore paw was raised in the 32-day fetus (D32, No. 13, F2) no further response to

stimulation of this area took place till the 39th day, when a slight fore-limb flexion was observed. At this same age (D39, No. 25, F1) stimulation led to a wink of the lids of the still closed eyes. The eyes open typically on the 56th day. Corrugation of the vibrissae pad, eye wink, pinna reflexes, mouth, tongue, lip, jaw movements, all occurred at approximately typical times. In comparison with the nostril the responses released were weak and of small amplitude. Local response, save for the localizing movements of the fore paws, was absent. In the later fetal ages this area became less, rather than more, sensitive. In late stages even a needle stimulus sometimes failed to elicit response when applied here in an otherwise active animal.

6, 7, 8, 9. *Side Face "A," "B," "C," "D"* (Summary only). In general, these areas remained insensitive longer than the ones considered above. In the case of the typical side face "B," save for a possible leg twitch at 37 days, the first real response occurred at 39 days (No. 27, F1). In the same way, for about a week before normal birth these areas again became relatively insensitive. In general, stimulation of these regions led to head and trunk movement, but not always so as to remove the head from the posture in which it was when stimulated. Limb movement was also released, but not nearly so frequently as in the case of areas 1, 2, 3, or 4. Corrugation of the vibrissae pad, eye wink, mouth movements, and pinna responses were also released. In general, the tendency to have any one of these responses released was greater if the spot stimulated happened to be near the areas of the response.

10. *Direct Touch Stimulation of the Conjunctiva of the Eye* (Summary only). In the observations recorded, the stimulus was applied directly to the naturally opened eye, and so not till the 56-day fetus (No. 47, F1) could the study be undertaken. Below, under number 73, the results of the photic stimulation of the eye are given. From the first fetus studied to the animal of approximately the age to be born two dominant responses were made. First, the eyeball moved in its orbit, and, second, the lids winked. As developmental age became greater, the variability in the wink became more marked. Thus, at 62 days a series of rapid winks are recorded (D62, No. 54) and a maintained "posture" of closure is recorded at 65 days (No. 58). The wink of the open eye should be interpreted

in view of the wink of the closed eye treated next below. Beside these responses there were some incidental trunk and limb movements, but on the whole this area seems set apart from the surrounding reflexogenous zones. Only one pinna reflex, for example, is recorded. The special nature of the nerve supply to this area must be considered in interpreting these results.

11. *Lower Eye Lid.* (D32, No. 13) Stimulation led to a twitch of the homolateral fore leg. With continued stimulation the same leg was brought up so that the paw touched with a brushing movement the side of the face. (D32, No. 12) The homolateral fore leg was extended after stimulation began and held in a posture with the limb up near the face as long as successive stimuli at the rate of about one a second were applied. (D32, No. 14, F2) Stimulation was followed by raising and adduction of the fore limb so that it seemed to come toward the point stimulated. (D32, No. 14, F3) N. R. (D34, No. 17) Stimulation was followed by a slight twitch of the head. The homolateral fore limb was also brought up toward the spot stimulated. A flexion at the carpus occurred after the limb was fully raised. (D35, No. 20, F2) The homolateral fore leg was brought up to the side of the face after stimulation. (D35, No. 19, F2) Stimulation led to a quick lateral flexion of the head, removing point stimulated from its previous locus. (D35, No. 19, F2) The homolateral fore paw came up and touched the eye as if to brush the stimulus away. (D35, No. 19, F3) Stimulation led to a quick lateral flexion of the head, accompanied also by a flexion of the rump. This movement was such as to bring the stimulus nearer the eyes rather than farther from it. (D35, No. 19, F3) Stimulation led to a contraction of the musculature of the closed lids and possibly of the eyeball itself. It seems as though the observed movement might possibly be not so much a winking movement as a movement of the muscles attached to the eyeball. (D36, No. 21) Stimulation led to a slight contraction of the closed lids with slight twitch of the homolateral fore limb. (D36, No. 22) N. R. (D37, No. 23, F1) The animal was "spontaneously active," but stimulation of the eye did not change this activity. (D37, No. 23, F2) A contraction of the closed lids followed stimulation. (D38, No. 24) "Winking" of the still unopened eye and apparent retraction of eyeball followed stimulation. There was also a twitch of the homolateral fore paw.

(D39, No. 27) Stimulation led to a sharp, quick wink of the closed lid with maintained contraction. General bodily activity accompanied this response. (D39, No. 26) Touching the closed eyelid elicited a movement which seemed to involve the eyeball inside as well as a sure contraction of the still unopened lids themselves. After stimulation the contraction of the orbicularis muscle was maintained for some seconds. (D39, No. 26, F3) Stimulation led to a wink of the lids of the as yet unopened eye. (D40, No. 28) A wink of the closed lids and a sharp lateral flexion of the head away from the stimulus followed stimulation. Both fore paws were then extended and abducted from the chest. (D41, No. 29) Stimulation led to a marked wink of the still closed lids and a twitch of the head. (D42, No. 30) Stimulation led to a wink, that is, a contraction, of the orbicularis muscle, together with a twitch of the pinna on the same side. After several seconds this was followed by an extensor thrust of all four paws and an extension of the neck, thus snapping the head up dorsally. (D43, No. 31) Repeated winking of the still closed eye followed stimulation. A slight extension of the fore legs was also noted. (D44, No. 32) To light touch there was no response, but to continued strong pressure a wink of the closed eye occurred. (D45, No. 33, F1) To a light touch no response was observed, but to heavier pressure a slow wink with somewhat maintained closure took place. (D45, No. 33, F2) A wink of the still closed eye followed stimulation. (D46, No. 34) An eye wink with somewhat maintained "posture" followed by several rapid winks followed stimulation. When the stimulation was repeated somewhat more strongly, the head was turned away and the homolateral fore leg brought up to the side of the face with digits of the paw flexed. The homolateral hind leg was also flexed and adducted to the side. In both of these cases a slight pinna reflex was noted. (D47, No. 35) Stimulation led to slight wink of the still unopened eye. (D47, No. 36) Stimulation was followed by a rapid flexion of the homolateral hind leg. The fore leg also moved toward the region of the stimulus. A repetition of the stimulus led to violent flexion of the neck and trunk and extension and adduction of the fore limbs so that the homolateral paw touched the point stimulated. (D48, No. 37) There was an immediate contraction of the musculature of the closed eyelids and possibly a movement of the pinna following con-

traction. (D49, No. 38) Contraction of the muscle of both the upper and lower lids and a slight extension of the head and head shake followed stimulation. (D49, No. 39, F4) Stimulation led to a contraction of the eyelids and rotation of the head away from the stimulus. (D50, No. 40) Eye wink of the closed eye, pinna reflex, and quick thrusts of both fore paws followed stimulation. (D51, No. 41) Eye wink of the closed eye followed stimulation. (D52, No. 42) Contraction of the still unopened lids, followed by an extension of the heterolateral and homolateral fore limbs, took place after stimulation. The homolateral limb was brought up almost to point stimulated. After this action the animal executed a series of complex movements apparently involving almost every muscle and joint of its body. (D53, No. 43) A wink of the still closed eye followed stimulation. (D54, No. 44) Stimulation led to several rhythmically repeated winks of the still closed lids. (D54, No. 45) Winking of the closed eye repeated several times after a single stimulus. (D55, No. 46) Stimulation led to a wink of the still unopened eye and a diffuse contraction of all facial muscles. The homolateral hind leg was also flexed and adducted so that the paw touched the shoulder. While in this posture it executed a series of rhythmic scratches. (D56, No. 47, F1) The eye is now open. Stimulation of the lower lid led to a closure of both lids with the closed posture maintained for a few seconds and then gradually opened. (D56, No. 47, F2) Stimulation led to an eye wink involving both lids. (D57, No. 48) Closing of eye, with greatest activity in lower lid, followed stimulation. (D58, No. 49) On stimulation both lids closed, the pinna responded, and a slight trembling of all four limbs was noticed. (D59, No. 50) To the lightest touch a series of winkings of the now open eye took place. The eye was maintained closed for a few seconds, then gradually opened. (D60, No. 51) Stimulation led to eye wink and maintained closure of the eye for several seconds. The stimulus was repeated with the same result, only this time the homolateral fore limb was brought up toward the face with flexion at elbow so as to execute a brushing movement on the face. (D61, No. 52) Eye wink to lightest touch took place. To a stronger stimulus, both fore legs were brought up to the side of the face and then flexed so as to brush it. The hind legs also twitched. (D61, No. 53) To the lightest touch a wink of slight

magnitude not quite closing the eye occurred. To a stronger stimulus there was a repeated "flutter" wink. (D62, No. 54) Stimulation led to eye wink. The fore and hind limbs then executed a series of coordinated swimming movements which were so effective that they removed the animal from the support on which it was resting under the surface of the bath. It was restrained only by the cord from swimming away. (D63, No. 55) Stimulation led to a tight closing of the eye, which was maintained for over a minute. (D64, No. 56) Sharp closure followed a single stimulus. Two winks, typical "palpebral reflexes," resulting from a single stimulus took place next. (D64, No. 57) Stimulation led to closing of the eye. (D65, No. 58) The stimulation led to an eye wink and then after about a second a vigorous shake of the head. This was followed by a violent response involving homolateral and heterolateral fore- and hind-limb trot movements. (D67, No. 60) Stimulation led to a rhythmic eye winking. A slight turning away of the head was also noted. The homolateral hind leg was flexed and adducted to the side and then executed a series of scratch responses. The homolateral fore leg was brought up to the neck.

12. *The Upper Eyelid* (Summary only). In general, the protocol given above under 11 also applies here. More pinna responses, however, were released.

13. *Brow* (Summary only). No response to stimulation resulted until the 40th day, when in a fetus (No. 28, F1) a wink of the still unopened eye and a pinna response occurred. These two responses were typical of all the responses elicited in this area. Some limb and trunk movement occurred, but taken as a whole this seemed to be incidental. One litter at 63 and two at 64 days gave no response at all to this form of stimulation. Although on the head and in a region that because of morphological prominence receives much random stimulation, this area seems relatively insensitive.

14. *Crown* (Summary only). This area is possibly about the least sensitive cutaneous zone of the organism. More than half of the forty fetuses which were stimulated in this region gave no response to the touch stimulus and some active fetuses did not even respond to needle stimulation when it was applied here. The typical response of those which were given was a quick twitch of the head. Palpebral and pinna reflexes were also elicited, as were, in

certain instances, limb and trunk movements. The difference between the poverty of responsive behavior released by this area and the richness of behavior elicited by the stimulation of the concha, next to be considered, is remarkable.

15. *Concha.* (For earlier response see page 357.) (D34, No. 17) A lateral flexion of the head away from the point stimulated followed the first touch. The homolateral hind leg was flexed and both fore legs were slightly twitched. Later the homolateral fore leg was brought up so that the paw on that leg brushed the side of the face. (D35, No. 20) A flexion of the lumbar segments of the back raised the rump toward the point stimulated. The head was rotated and the homolateral hind leg flexed and adducted to the belly. The homolateral fore leg was adducted to the neck. With repeated stimulation there was a repetition of the movements just noted. (D36, No. 22) Stimulation led to a twitch of the pinna and an extension of the neck, serving to produce an upward thrust of the head. (D37, No. 23, F1) A slight lateral flexion of the trunk correlated with a twitch of both fore paws followed stimulation. On a second stimulation there was a slight rotation of the head to one side and a sharp jerk of both fore paws. The result of this response was a movement of the ear away from the posture it had been in when it was stimulated. (D37, No. 23, F2) Lateral flexion of the neck and trunk, thus moving the ear away from the stimulus, occurred as a result of a light touch. (D38, No. 24) Stimulation led to a slight lateral trunk flexion which was accompanied by a raising of the rump on the homolateral side. Both fore legs twitched and the homolateral fore limb was carried across the chest to the heterolateral side. There was no indication of "localization" in these limb movements. When stimulated again in the same spot the only response was a lateral shake of the head. (D39, No. 27) Definite elevation and depression of the jaw, lateral head shake, violent flexion and extension of all four legs followed stimulation. Lateral sigmoid flexure of the trunk also took place, after which the homolateral hind leg was flexed and adducted to the side. When the hair stimulus was held persistently in the concha the animal squirmed and turned both body and head, as if to rid itself of the stimulus. (D39, No. 26) Stimulation led to a contraction of the pinna as a whole. At a later time this response was also observed to occur

"spontaneously." (D39, No. 25) After stimulation the left hind leg was flexed and adducted toward the side. The fore leg was also extended and raised part way, but not so as to touch the spot stimulated. (D40, No. 28) With very slight stimulation a pinna twitch was observed. With rather strong stimulation the head was rotated toward the stimulus and the homolateral hind leg flexed and adducted to the body. Both fore legs were extended, but the homolateral limb was also elevated and adducted so that its paw touched the face. A lateral flexion of the trunk and a movement of the mandible were also recorded. (D41, No. 29) Stimulation led to a slight lateral shake of the head. The neck was flexed and then the dorsal flexure passed down the whole back of the fetus. A sharp twitch of the homolateral hind leg also took place. (D43, No. 31) Stimulation led to a quick rotation of the head away from the stimulus, a contraction of the pinna, a wink of the eyelids of the still unopened eye, and a series of rapid paddling movements of both hind and fore legs. The trunk also was laterally flexed. (D44, No. 32) Light touch led to the contraction of pinna only. But continued strong stimulation elicited a slight writhing movement of the trunk with a bending of the rump in direction of stimulus. There was also a slight flexion of the homolateral hind leg. (D45, No. 33, F1) N. R. (D45, No. 33, F2) A slight extension of the neck led to an upward toss of the head. At the same time the two hind legs were extended away from the body and the trunk flexed so as to remove the head from the position in which it had been when stimulated. (D46, No. 34) No response to a single touch, but when the stimulus was maintained inside the concha for some time a lateral flexion of the lumbar segments of the back raised the rump laterally. This flexure then became even more marked as the homolateral hind leg was flexed and adducted toward the ear. The neck was also flexed so that head, trunk, and rump made a perfect "C." The fore legs were extended and adducted to the head, but did not touch the ear at first, but with continued stimulation the paw of the homolateral fore leg did brush the face. At this time the limb was flexed at the carpus and digits. At a later time to the same stimulus the homolateral hind leg was brought up sharply and executed a series of rhythmic scratch movements which did not, however, touch any part of the animal. (D47, No. 35) Stimulation led to a contraction

of the lids of the still unopened eye and a corrugation of the muscles under the pad of the vibrissae. (D47, No. 36) Stimulation led to a twitch of the hind leg and also a contraction of the pinna. Lateral flexion of the trunk and elaborate paddling movements of both fore and hind legs followed stimulation. (D48, No. 37) A slight pinna twitch followed first stimulation. Then a sudden unilateral bending of the rump took place which raised the rump and hind legs toward the surface of the water and away from the supporting stand. (D49, No. 38) No response to light touch, but to continued stimulation first the muscles of the ear contracted and then the homolateral fore limb was extended and adducted with the digits of the paw of the limb held in a maintained flexure. The paw was thus mechanically "hooked" behind the ear where it remained for several seconds. The paw then slipped down, actually forcing the stimulus hair away from the point which it was touching. During this procedure the homolateral hind limb was adducted and extended toward the point stimulated. Save for the fact that it became mechanically caught in the cord, it would presumably also have been adducted so as to touch the point stimulated. (D49, No. 39) Stimulation was followed by an extension and adduction of the homolateral fore limb so that later the paw of this limb brushed past the point stimulated. There was also much general activity of the fetus during this response. (D50, No. 40) The head was rotated slightly after stimulation and then with a minimum of other body movements the homolateral leg was brought up so that the digits of the paw cupped (i.e., flexed) over the stimulated concha. The paw then brushed on down the face and back to its resting place. After an interval of rest this stimulus was repeated. A shake of the head, a pinna reflex, and a flexion of both hind limbs took place. In this flexion the homolateral limb was more involved than the heterolateral limb. An extensor thrust out into the bath and away from the body also took place in both fore legs. (D51, No. 41) Stimulation led to a pinna contraction and to a shake of the head. The stimulus was repeated and the homolateral fore leg was brought up to the base of the ear. Later the pinna contracted and the homolateral hind leg was flexed and adducted; it then executed a series of rhythmic scratches. The head was also given a quick lateral shake. (D52, No. 42) Lateral flexion of the neck and trunk took place after stimulation. The

homolateral hind limb was sharply flexed and the foot became caught in the umbilical cord. This limb, however, executed a series of rhythmic "scratch responses" when held in this posture. Soon the homolateral fore limb was adducted to the point stimulated and then the homolateral hind limb, now free, actually touched the point stimulated. When the hind limb was held by the experimenter so that it could not touch the point stimulated the homolateral fore limb was adducted so that its paw touched the stimulated region. (D53, No. 43) A combined lateral and dorsal flexion of the neck "twisted" the head away from the point stimulated. A strong extension of the whole back then took place which was followed by alternate paddling of the fore legs and by a flexion and adduction of both hind legs to the belly. After this the whole trunk continued a series of writhing flexures. (D54, No. 44) Response of the muscles of the pinna followed stimulation. Upon more violent stimulation there was a slight adduction of the homolateral fore limb to the head. (D54, No. 45) Stimulation led to local contraction of the musculature of the side of the face. (D55, No. 46) Stimulation was followed by a lateral shake of the head and by a marked flexion of the whole back. The homolateral hind limb was now adducted and elevated so that the nails of the digits actually entered the concha, all the while executing scratch responses. After relaxation the homolateral fore paw was raised to the same point and executed brushing movements. It must be obvious that in order to bring about the successful scratch of the ear by the hind leg the lateral flexure of the trunk into a "C" form had to be maintained. (D56, No. 47, F1) Stimulation led to closing the now open eye, lateral flexure of the head away from the stimulus, a pinna contraction, and an upward and outward thrust of the homolateral fore leg. After an interval of rest the stimulus was again applied and similar responses recorded. (D56, No. 47, F2) A light touch led to a contraction of the lower part of the pinna. When the stimulus was given again the eye winked. When the pressure was made greater the head was rotated toward the stimulus and both fore legs were extended. The homolateral hind leg was flexed and adducted to the shoulder, where it executed a rhythmic scratch movement. This movement brought the digits of the paw almost but not quite to the point stimulated. (D57, No. 48) A slight touch led to a local contraction of the

muscles about the point stimulated. This response was followed by a toss of the head backward. On continued stimulation the neck was extended and the head thrust back. The ear muscles sharply contracted and lateral and dorsal flexions of the trunk took place. The fore paw was then lifted toward the point of stimulation and the hind limb flexed and adducted so as almost to touch the point stimulated. (D58, No. 49) Stimulation led to a pinna response and eyeball movements. (D59, No. 50) To a very light touch stimulation a contraction of the pinna and the musculature surrounding the concha took place. The eye also winked. A flapping of the opposite pinna was also observed to be of greater amplitude than the reaction of the one stimulated. On second stimulation the head was shaken laterally as if to toss off the stimulus. On stronger stimulation the head was again shaken and the heterolateral hind leg was adducted to the side and the homolateral fore leg extended and adducted toward the ear. When the stimulus stopped the leg relaxed. On continuing the stimulus the head twisted and bent down, and the homolateral fore leg was raised so that it touched the ear. At the same time the homolateral hind leg was flexed and adducted and brought up near to but not touching the ear. Eye wink and pinna twitches occurred frequently and the limbs were flexed at all joints. (D60, No. 51) Stimulation led to a flexion of the whole back and especially of the neck, so that the head was thrust down. The rump was then pulled up laterally and the homolateral hind leg brought up so that the nails of the digits rhythmically scratched the concha. (D61, No. 52) A lateral flexion of the lumbar segments of the back raised the rump and a similar flexure in the neck produced a true lateral "C" posture after stimulation. The homolateral hind leg was then adducted to the head and the digits so flexed that the inside of the ear was thoroughly scratched. After this posture had relaxed, the homolateral fore leg was brought up and brushed by the ear. (D61, No. 53) Stimulation led to several twitches of the pinna, a shake of the head, and a forward jerk of the body. At the same time the homolateral hind paw was brought up toward the ear and later the same movement was executed by the homolateral fore paw. (D62, No. 54) Vigorous, repeated, flutter-like shake of head followed stimulation. After this response rhythmic "swimming" movements took place involving all four limbs.

It was impossible to tell whether or not all limbs were synchronized, and unfortunately it was not photographed. (D63, No. 55) A sharp shaking of the head followed stimulation. One of the eyes which had been closed was opened and the homolateral fore paw was raised to the face. The pinna twitched and then the fore limb was again brought up, but this time it accurately localized the place stimulated. The homolateral hind leg executed a continued "scratch response" during the last movement. (D64, No. 56, F2) Stimulation led to a slight twitch of the head. This was followed by a movement of the homolateral limb such that by means of articulation at shoulder, elbows, wrist, and digits the digit tips moved over the concha with a brushing movement and then back to the original posture. Tremor-like reflexes of pinna took place frequently. (D64, No. 57) Local responses of the ear were observed to follow stimulation. Then both fore legs were sharply flexed and the abdominal wall retracted convulsively. (D65, No. 58) To a very light touch a localized pinna reflex took place. With a somewhat stronger and maintained pressure the neck flexed, thus dipping the head toward the chest, and both fore legs were flexed to the belly at first, but then the homolateral hind leg was brought up to the head, where it scratched the ear, the head being flexed laterally, as it were, to meet it. With continued strong stimulation the homolateral fore limb was raised to the face, where it executed brushing movements, but did not touch the ear. (D67, No. 60) With a single touch, the head rotated away sharply, the eye winked, a pinna response took place, and the homolateral hind leg was brought up slightly. With continued stimulation, however, the homolateral hind leg was flexed, adducted, and finally elevated up over the shoulder, and the head so turned in the direction of the paw that a rhythmic scratch of the ear was executed.

16, 17. *Pinna "A" and Pinna "B"* (Summary only). Stimulation of both these zones led to very similar results. The first response was in a 32-day fetus and involved a lateral flexion of the trunk. As fetal age increased, stimulation released limb and trunk movements. Some of the limb responses definitely touched the spot that had been stimulated. From the 35th day on to normal birth the characteristic response to stimuli applied in either of these areas was a contraction of the whole pinna or, in certain instances, of the

pinna for a few mm. around the locus of the point touched. Eye wink was also elicited by stimulation of these areas in a number of cases. These areas are interesting in that they demonstrate the presence of a definite reflex and also in the fact that, while the term "pinna reflex" or Preyer's reflex is used extensively in the literature, it is itself subject to variation. Careful analysis would probably show that in hardly any two instances are the actual "responses" of this supposedly unitary reflex identical. The most delicate motor response that the present writer has ever seen in a guinea-pig fetus was a response of Fetus 1 of Animal 48 at 57 days to stimulation at spot "B." Here only the faintest flicker of a response occurred in the extreme outer tip of the pinna, but time relations indicated that this was undoubtedly a response involving the central nervous system and not a direct muscle response.

18. *Neck Ventral* (Summary only). Stimulation of this zone led to diverse responses, of which the most typical seems to have been the flexion of the neck so that the head was bowed toward the chest and the lifting of the homolateral fore leg. After the 43rd day local contraction of the musculature underlying the point stimulated occurred. This may be called a panniculus carnosus reflex. In the last week before birth there were a number of fetuses in which neither pressure nor needle stimulation at this spot elicited response.

19. *Neck Dorsal.* (Summary only). This area was relatively insensitive, as judged by the vigor of the movements which it elicited. At first on stimulation this area released trunk and especially neck responses. The latter were probably dominantly extensions. Local skin contraction was also almost a constant element after the 39th day. In a number of instances stimulation of this spot released pinna reflexes and eye winks. In general, this area does not seem to have the same "significance" for the behavior of the animal that is possessed by, for example, certain of the head zones.

20. *Side above Shoulder* (Summary only). At 34 days the first response recorded for this area was an elaborate movement of the homolateral fore limb which involved a movement of the shoulder girdle. At 39 days a local muscle response was noted as occurring just below the point stimulated. From this day on, this skin response became something of a constant. There can be no doubt that this

point of stimulation led to more shoulder movement than had been seen before. Incidental head reflexes, hind-limb scratch reflexes, and general trunk responses were also elicited.

21. *Back above Shoulder* (Summary only). The general pattern of response elicited from this zone was similar to that of number 20 just considered. This region was less sensitive, however, as judged by behavior. In fetus after fetus the only response elicited by stimulation here was a local skin corrugation or a shudder and a twitch of the trunk. In certain instances, however, definite limb and head reflexes were released from this area.

22, 23. *Side Back "A" and "B"* (Summary only). These two areas, as well as number 24 which is considered below in full, seemed to release very similar responses. In both areas trunk movements as well as occasional limb flexions were elicited. The most characteristic responses of the area, however, were the local skin reflexes which were very regular and well localized. In general, these areas may be characterized as relatively insensitive.

24. *Side Back "C."* (D32, No. 13) N. R. (D34, No. 17) N. R. (D36, No. 22) N. R. (D38, No. 24) A quick lateral flexion of the rump, accompanied by a twitch of both hind and fore limbs followed stimulation. Later the fore limbs were extended and adducted so that they touched the face. (D39, No. 27) A skin twitch, which rippled caudally, followed stimulation. (D39, No. 26, F4) Stimulation elicited a lateral flexure of the lumbar segments of the spine away from the stand which supported the fetus. (D40, No. 28) Stimulation led to a skin twitch near the point of the touch. (D41, No. 29) Local skin contraction and raising of rump followed lateral flexure of the lumbar spine. (D42, No. 30) Stimulation led to a local skin contraction which spread by ripples of corrugation almost to the head. A lateral bending of the lumbar segments of the spine also took place. (D43, No. 31) Stimulation was followed by a slight unilateral bending of the lumbar spinal segments, thus lifting the rump away from the supporting stand. The skin was also corrugated at the point of stimulation. (D44, No. 32) There was a unilateral bending of the lumbar spine and wrinkling of the skin following stimulation. The heterolateral leg was extended and there was a slight flexion in the homolateral hind leg. The same response was given to a more intense stimulus.

(D45, No. 33, F2) A light touch led to a slight flexion of the rump so as to lift the fetus up from the supporting stand. The skin was also corrugated near the point of stimulation. (D46, No. 34) A local corrugation followed by rippling of the skin in a cephalic direction occurred. (D48, No. 37) A skin response took place at the point stimulated. (D49, No. 38) The thigh was so flexed after stimulation as to change the shape of the musculature immediately beneath the point stimulated. A toss of the head and a jerk of the homolateral fore limb also took place. (D50, No. 40) The homolateral hind leg was flexed at the hip and the knee following stimulation. A general flexion of trunk muscles then took place, bowing the animal together. (D51, No. 41) With continued stimulation the hind legs were flexed and adducted to the belly. A slight lateral flexion of the lumbar region of the back and a movement of the fore legs and head followed stimulation. (D52, No. 42, F1) No response occurred to light pressure; but to deep pressure a lateral flexion of the trunk was recorded with bending up of the lumbar region away from the support. A slight extension of the homolateral fore and hind limbs also took place. (D52, No. 42, F2) A stimulus similar to that just recorded led to exactly the same response as that recorded for Fetus 1. (D53, No. 43) Stimulation led to a local twitch of the skin when the stimulus hair was brushed over this point. There was no response to single-touch stimulation. (D54, No. 44) A single touch led to no response. Stroking stimulus led to an extension of the spinal segments in the lumbar region. (D54, No. 45) N. R. (D55, No. 46) A slight twitch of all four paws followed stimulation. (D56, No. 47, F1) N. R. (D56, No. 47, F2) A strong pressure stimulus led to a strong backward thrust of both hind legs and an upward stretching movement of the anterior part of the body. There was no response to a slight touch. (D57, No. 48) N. R. (D58, No. 49) A twitch of the leg on the same side followed by a slight flexion of the leg on the other side followed stimulation. (D59, No. 50) No response took place to a touch or to stroking, but to deep pressure a strong outward kick of both hind legs took place. A slight upward extension of the head and twitching movements of both fore legs also occurred to this stimulus. (D60, No. 51) No response could be elicited to pressure or to a needle, although the animal was otherwise active. (D61,

No. 52) No response was given to touch or pressure. To vigorous needle stimulation the fore legs were slightly extended and a flutter of the flexed hind legs took place. Rump twisting then occurred and the homolateral hind leg executed a scratch response. (D62, No. 54) Stimulation led to hind-leg extension and a slight body extension. (D63, No. 55) N. R. (D64, No. 56) N. R. (D65, No. 58) N. R. (D67, No. 60) N. R.

25. *Side Back "D"* (Summary only). This area is very similar in the responses which it releases and in its slow rise of sensitivity and its loss of sensitivity to area 24. Here, however, local skin responses during the active period were noted. In one fetus of 57 days (No. 48) stimulation led to a rapid, rhythmic series of premature breathing movements (Ahlfeld's movements). During the last week of fetal life this area was, with minor exceptions, absolutely insensitive to touch or needle stimulation.

26. *Rump* (Summary only). The responses which were released when this area was stimulated were different from those elicited by the four adjacent areas just considered. This area, when stimulated, was more effective in calling out responses. Stimulation here led to many trunk movements, most of which were extensions. Limb movements, predominantly but not exclusively of the hind legs, were also commonly elicited. No local skin corrugations were observed in this area.

27. *Hip* (Summary only). Stimulation of this region was effective in bringing about response in a 32-day fetus (No. 13). This response included a flexion of the hind limbs which proved to be the most typical response to stimulation in this area. It was also the dominant response at 65 days (No. 58). Trunk movements and even head reflexes were also elicited by stimulation of this area, but it was apparently dominantly concerned with the hind limbs and especially with flexion of the homolateral hind limb.

28. *Anus Area.* (D34, No. 17) N. R. (D35, No. 20) Stimulation was followed by a slight movement of fore and hind legs. (D36, No. 21, F2) Stimulation was followed by a drawing in of rump with an extension of both hind legs and a pulling together of the thighs so as to "protect" the region stimulated. (D36, No. 22) N. R. (D37, No. 23, F1) Stimulation led to a slight jerk of the rump followed by rather general body and limb movements. (D37, No.

23, F2) N. R. (D38, No. 24) A retraction of the muscles of the anus area with an extension of both hind limbs and a flexion of the rump so that it curved in toward the belly occurred. (D40, No. 28) The anal musculature was sharply constricted, the rump curved in, and the hind legs flexed and adducted to the chest. The fore legs were then flexed, so that fore and hind legs came in contact. (D41, No. 29) A quick and violent retraction of rump so as to pull the anal opening in between the hind legs followed stimulation. (D42, No. 30) Stimulating the anal region elicited an expansion and then a constriction of the muscles immediately surrounding the anus. (D43, No. 31) Stimulation led to a sudden contraction of the anal musculature inward toward the umbilical cord. A slight movement of the hind legs also took place. (D44, No. 32) To stimulation the muscles around the anus contracted and relaxed, producing first an opening and then a closing. Stronger stimulation led to both hind legs being strongly flexed and adducted so that the thighs came to cover the area of stimulation. (D45, No. 33, F2) A sharp extension of both hind legs and a slight twitch of the anal musculature followed stimulation. (D46, No. 34) Stimulation led to the homolateral hind leg being pushed against the ventral abdominal musculature and maintained there. The rump and anal musculature then "curved in." The head was also flexed laterally in the direction of the stimulus and the homolateral fore leg brought down so that it touched the hip. On restimulation a palpitating contraction of the anal musculature took place, accompanied by an extensor thrust of both hind legs. (D47, No. 35) Stimulation led to a long series of movements, in which much of the body cooperated and in which both "heels" were brought up toward the point of stimulation. (D48, No. 37) A quick sharp contraction of the musculature around the anal region followed stimulation. A curving in of the rump and flexion of the hind legs also took place. (D49, No. 38) Stimulation led to the rump's being pulled in, the hips thrust back and pulled together, thus quite effectively protecting the area from further stimulation. The fore limbs and neck were also extended. (D50, No. 40) Stimulation led to the body's being contracted together until it seemed like a sphere. The rump was then pulled in under the hind legs; the homolateral hind leg next contracted and rotated laterally, so as to bring the sole of the

foot against the point stimulated. (D51, No. 41) A light touch led to rhythmic pulsating contractions of the surrounding musculature. To continued stimulation a dorsal flexion of both hind legs and curving in of the rump took place. (D52, No. 42, F1) No response to touch stimulus. To continued rather strong stimulation, however, a twitch of both hind legs and especially of the right hind leg took place. This was followed by a jerk of the body and by local contraction of the musculature about the anus. (D52, No. 42, F2) To light pressure a pronounced contraction of muscles around the anus took place. The musculature of the rump was then "pulled in." (D53, No. 43) Stimulation led to an active constriction and then dilation of the muscles surrounding the anus. (D54, No. 44) Touching led to no response. Continued stroking, however, was followed by pulling in of the anal area with a flexion and adduction of both hind legs to the body. (D54, No. 45) Stimulation led to a simultaneous backward thrust of both hind legs and a contraction of the musculature in the region stimulated. (D55, No. 46) Local contraction of the muscles around anus followed stimulation. Contraction of the general rump musculature and pulling in of the rump between the legs with clamping together of the heels and feet of both hind limbs so as to "protect" the area also followed stimulation. The homolateral fore limb then extended, the neck muscles also extended, and the head was raised and maintained in the posture noted. (D56, No. 47, F1) Contraction of the anal musculature followed stimulation. A slight backward thrust of both hind legs followed. On repetition the musculature of the anal area contracted, but the limbs did not move. (D56, No. 47, F2) Contraction of local musculature followed stimulation. A slight flexion of the homolateral hind leg also occurred. (D57, No. 48) A local contraction of the muscles around the anus was followed by a pulling together of the hips and curving in of rump so as to "protect" the area. This movement was accompanied by slight lateral bending of the rump away from the supporting stand. (D58, No. 49) Stimulation led to a contraction and pulling in of the muscles surrounding the anus. At the same time the rump bent inward. The spine was then flexed, both fore legs were flexed, and the hind legs extended so that the paws almost touched. A vigorous stretch of both hind legs and even more extreme flexion of the neck then took

place. (D59, No. 50) Stimulation led to a violent outward thrust of both hind legs and a change in the muscles of the anal region. Repeated stimulation led to the same response. A very light touch led to an opening of the anal area without any correlated hind-limb kick. (D60, No. 51) Stimulation led to local muscle contraction which was followed by a slight extension of the homolateral fore leg. A needle stimulus led to a more vigorous response which included alternate paddling movements of the hind limbs, a curving in of the rump, and a slight extension of both fore limbs. (D61, No. 52) Stimulation led to a contraction of the anal musculature. A needle stimulus led to a sharp flexion and bringing together of both hind legs as if to protect the region stimulated. (D61, No. 53) Stimulation was followed by an extensor thrust alternately of both hind legs. A twitch of the muscles about the anus also took place. (D62, No. 54) Local contraction of the musculature followed stimulation. Then there was a sharp synchronized backward thrust of both hind legs and alternate paddling of the fore legs and an upward thrust of the head occurred. (D63, No. 55) A quick localized abduction and then adduction of the homolateral hind leg took place after stimulation. (D65, No. 58) Stimulation led to a contraction of muscles around the anus and an accompanying extrusion of the orifice of the anus. (D67, No. 60) Stimulation led to a slight twitch of the muscles about the anus.

29. *Knee* (Summary only). At first this area was relatively insensitive, as judged by the behavior released from stimulating it. Later it led to movement of the stimulated limb, sometimes extension, sometimes flexion. In certain instances the digits of the limb were either flexed or extended, usually in harmony with the dominant movement of the limb. In several instances local skin twitches resulted from stimulation.

30. *Leg* (Summary only). This was a relatively insensitive area. It did not become able to release behavior until 35 days and again lost it some time before birth. The dominant movements were of the stimulated limb and of the heterolateral hind limb. Like area 29 stimulation here also released some local skin responses.

31, 32. *Foot "A" and "B"* (Summary only). In both of these cases the stimulus was applied to the dorsal aspect of the foot. In many fetuses at all of the ages studied stimulating these spots led to

no response. A flexion of the stimulated leg was the most common of the responses that did occur, although occasionally an extension of the digits was released. The poverty of the responses elicited by this stimulation is in marked contrast to that noted under 65 below, when the plantar surface, only a few mm. away, is stimulated.

33, 34, 35. *Hind Foot Toe "A," "B," and "C"* (Summary only). In general, these areas were insensitive to stimulation, as judged by behavior, until the 36th day. Even after this in many instances stimulation led to no response. In general, the most common response to any of these stimuli was an extension of the whole stimulated limb or the ankle of the limb alone. In all but one instance, when the digits or one digit responded, the response was an extension and not a flexion.

36. *Abdomen Side "A"* (Summary only). Beginning at the 34th day (Animal 17) and on through most of the gestation period the typical response elicited by the stimulation of this area was a flexion and an adduction to the stimulated region of the homolateral hind leg. In a number of cases the typical "scratch response" was released in the same leg. Corrugation of the skin under the place stimulated was also an almost uniform response. No responses were elicited between the 63rd and the 67th day.

37. *Abdomen Side "B."* (D32, No. 12) Stimulation led to a slight contraction of the muscles of the body wall. (D32, No. 14) After stimulation the homolateral fore paw was flexed and moved in the general direction of the stimulus. (D34, No. 17) Stimulation led to a slight flexion of the fore leg with pronounced flexion of the digits of the paw of that leg. (D35, No. 20) N. R. (D35, No. 19) Stimulation led to a quick extension of both fore limbs with a flexion at the elbows. After this movement the homolateral fore leg was moved toward the point of stimulation. (D36, No. 21, F2) No response to touch stimulus. On continued stroking first the heterolateral fore limb and then the homolateral fore limb were adducted toward but not touching the point stimulated. (D38, No. 24) No response to touch stimulus. On continued stroking stimulation, however, the fore limbs were raised toward the head, the head flexed slightly, and the lumbar spine laterally flexed to the beginning of a "C" movement. After this response a more violent contraction of apparently almost all of the trunk muscles and a flexion of all

four limbs took place. (D40, No. 28) Stimulation led to a corrugation of the skin which spread by a series of ripples in a cephalic direction. (D41, No. 29) No response took place to touch stimulation at first, but a second stimulus led to the hind limb's being flexed and adducted to the point stimulated. Both fore paws were then brought up and brushed down both sides of the snout. (D42, No. 30) Stimulation was followed by a contraction of the skin directly underlying the spot stimulated. At the same time a downward thrust of the left fore leg took place, which was followed by an extension of the neck and consequently by a backward toss of the head. Repeating the same stimulus elicited again a local contraction of the skin, together with a slight flexion of the homolateral hind leg so that it almost touched the point stimulated. (D43, No. 31) The animal was continuously active and stimulation seemed merely to increase the violence of the trunk flexure and limb movements which were occurring. (D44, No. 32) To light touch a slight twitch of both hind legs took place. To a strong stroking stimulus, however, both fore legs were brought down to the chest, while the hind legs were flexed and adducted to the belly. (D45, No. 33, F2) A slight pulling in of the abdominal wall and a downward twitch of the homolateral fore leg followed stimulation. A slight flexion of both hind legs also took place. (D46, No. 34) N. R. (D48, No. 37) A slight contraction of the underlying musculature, a quick extension of the neck, and consequent snapping of the head upward followed stimulation. The fore legs were then adducted to the face and both hind legs extended backward away from the body. (D49, No. 38) A bilateral flexion progressing from head to tail and involving flexion and extension of all four legs followed stimulation. (D49, No. 39) Stimulation elicited a sudden flexion and adduction in the homolateral hind leg so that it touched the abdomen with a slight sweeping or pushing movement at the point of stimulation. This stimulus was given again and exactly the same thing occurred. After the sweeping or pushing movement passed the point of stimulation the leg was then extended. (D50, No. 40) Stimulation led to a localized "sucking in" of the abdominal musculature. (D51, No. 41) A retraction of the musculature at the point stimulated and a twitch of all four legs followed stimulation. (D52, No. 42, F2) Stimulation led to a contraction of the abdominal mus-

culature which was deeper than the mere panniculus carnosa reflex. (D53, No. 43) The heterolateral fore paw was flexed after stimulation. (D54, No. 44, F1) N. R. (D54, No. 44, F2) Continued stimulation led to a rigid drawing down of both fore legs toward the stimulus. (D54, No. 45) N. R. (D55, No. 46) No response to a single touch stimulus. On continued stimulation a slight jerk of all limbs and trunk took place. (D56, No. 47, F2) N. R. (D58, No. 49) N. R. (D59, No. 50) N. R. (D60, No 51) No response to any form of pressure stimulus. A slight touch of the needle led to timed responses of extension and flexion of the homolateral fore limb involving separate articulation of each joint. After this the hind legs began to execute a rhythmic beat. (D61, No. 52) No response to touch. To a needle stimulus, however, the homolateral fore leg flexed slightly toward the chest and both hind legs twitched. (D62, No. 54) Stimulation led to a flexion of the neck and consequent bowing of the head followed by an extension of the whole trunk. All four legs were also slightly extended. (D63, No. 55) N. R. (D65, No. 58) N. R. (D67, No. 60) N. R.

38. *Shoulder* (Summary only). At 34 days a flexion of the stimulated limb was released. This response continued to appear as the most constant factor in the protocols up to the 54th day, after which with one exception no response could be elicited from stimulating this area. Local contraction of the skin was recorded in a number of fetuses.

39. *Elbow* (Summary only). Stimulating this area led to a flexion of the stimulated limb at 34 days (No. 17). After this until the 51st day this was the typical response, although some special features appeared from time to time, such as the scratching of the stimulated spot by the heterolateral hind leg. After the 51st day, although there were one or two apparently sporadic responses, this area became apparently insensitive.

40. *Fore Arm* (Summary only). The first response to stimulating this area occurred at 35 days (No. 19). It involved a flexion of the stimulated limb. At later ages this same response was common, although extension also appeared. It seems possible that stimulation released "movement" and that "extension" or "flexion" depended upon the position of the limb or some other secondary determinants operating at the time of stimulation, as no special temporal order

can be discovered in these responses. After the 49th day this area became virtually insensitive to touch stimulation, and even needle stimulation often failed to elicit any response in an otherwise active fetus.

41. *Wrist* (Summary only). Stimulating the dorsal aspect of the carpus led to a flexion of the whole stimulated limb at 34 and 35 days. After this time most of the responses also included either marked extension or flexion of the carpus and the digits of the paw. There is no tendency apparent in the choice between flexion and extension.

42, 43, 44, 45. *Dorsal Aspects of the Front Paw Toes "A," "B," "C," and "D."* (Summary only). The first response to stimulation of this area was a flexion of the whole limb. Throughout the whole course of fetal life flexion or extension of the whole limb continued to occur. But the characteristic response was a local movement of the toes; in 21 cases this was recorded as an extension, in 5 cases as a flexion. In one case at 46 days (No. 34) flexion at the carpus and flexion of the digits followed stimulation. After this, extension of the digits occurred and then an alternation between "fanning" (extension) and "cupping" (flexion) of the digits was several times repeated. In a number of cases after 50 days the single toe upon stimulation was extended.

46. *Dorsal Mid-line of the Snout* (Summary only). The typical response to stimulating this area seemed to be head extension, although once or twice, including once at the 36th day (No. 21), head flexion occurred. In many instances the fore limbs were also brought up toward or actually touching the point stimulated. Special head reflexes, including lip, tongue, jaw, eye, and pinna responses were elicited at times. The hind-leg scratch reflex was also released in several fetal stages.

47. *Dorsal Mid-line of the Crown* (Summary only). This spot was in general not sensitive. At first when stimulation of this spot did release movement, that movement was of the head and trunk. By the 40th day (No. 28) an eye wink appeared and this was then an almost constant response. Pinna responses were also common. Other incidental movements occurred to stimulation of this zone, including a hind-leg scratch reflex and fore-limb extension.

48. *Dorsal Mid-line of the Neck* (Summary only). The first

responses in this area were twitches of the fore and hind legs and flexions of the trunk. In many instances the fore legs were raised toward the point stimulated. From the 39th day on, corrugation of the skin under the place stimulated was almost a constant response. After the 54th day the responses to stimulation in this area became few. During the mid course of the active period, 39-54 days, head reflexes and scratch reflexes were also elicited.

49. *Dorsal Mid-line Shoulder* (Summary only). At 32 days (No. 13) stimulation here led to rotation of the head, "paddling movements" of all four legs, and a writhing of the trunk. In this response were contained all of the typical elements of the later responses made by older fetuses to stimulation of this area, except local subcutaneous muscle responses, which occurred first at 37 days (No. 23) and then, quite regularly, up to 58 days (No. 49). After this period there was practically no response to pressure or needle stimulation in this area.

50. *Dorsal Mid-line Back* (Summary only). The first sure response to stimulation in this area was a slight local skin-muscle twitch (D38, No. 24) which also involved special head and limb movements. Apparently incidental limb responses were elicited after this at times, but the one constant response was the local skin contraction. From the 56th day to birth no responses could be elicited, save that in No. 54 at 62 days a local skin response followed stimulation.

51. *Dorsal Mid-line Hip* (Summary only). From the 38th to the 52nd days stimulation in this area led to trunk movement, dominantly, but not uniformly, extension. During this period, also, local skin responses were very regularly elicited. After the 53rd day practically no response could be elicited, even by needle stimulation.

52. *Dorsal Mid-line Rump* (Summary only). At 38 days (No. 24) a flexion of the rump occurred, such that it was pulled in between the hips of the hind legs. This was the first response elicited by stimulating this region and remained the characteristic response up to the 48th day, when characteristic responses ceased altogether, save for an incidental twitch in limb or trunk. At 67 days, however, (No. 60) the hind legs were flexed up to the belly and a flutter response of the paw of the hind limb nearest the surface took place. This was repeated once.

53. *Ventral Mid-line Snout* (Summary only). On the 36th day (No. 22) that is, at the first recorded response, and at the 67th day (No. 60), that is, on the day before birth, stimulation of this area led to the fore leg's being brought up and touching the face. Between these two extremes many other responses were elicited, the most characteristic of which was an extension of the neck and consequent elevation of the head. Various head reflexes, including responses of one side of the upper lip alone, tongue movements, jaw movements, eye winks, and pinna responses, were all elicited. During the last ten days of the gestation period this area was not uniformly sensitive.

54. *Lip Groove* (Summary only). At 35 days (No. 19) stimulation led to a quick extension and withdrawal of the head and a slight movement of both fore legs. A slow movement of the right paw toward the snout then occurred. This movement resulted in a brushing or waving movement of both fore limbs, ending with the *right paw's touching the left side of the face.* On the following day (No. 22) mouth and lip movements were observed. These two patterns of response seemed to be the dominant ones released by stimulating this area. As the gestation period advanced, stimulation of this area at one time or another released all of the typical head reflexes, including those of the tongue, jaw, eye, and pinna.

55. *Upper Gum.*[4] (D32, No. 12) Stimulation of the lips led to a bringing up of both fore paws as if to brush away the offending stimulation. (D34, No. 17) Movement of the left fore leg, brushing by but not touching face, followed stimulation. (D36, No. 22) Turning of the head toward left and fore-leg and hind-leg twitches followed stimulation. (D38, No. 24) A localized contraction of lip on the side stimulated followed by rhythmic tongue movements resulted from stimulation. (D40, No. 28) Contraction of the upper lip, retraction and movement of the tongue, and an upward pushing of both fore limbs followed stimulation. (D41, No. 29) Stimulation led to a wider opening of the mouth and to tongue protrusion. (D42, No. 30) A slight lip contraction followed stimulation. (D43, No. 31) The tongue was pulled back into the mouth, and jaw

[4]In many cases, because of lack of dexterity on the part of the experimenter, this point was really "lip and gum" rather than "gum" alone.

movements followed stimulation. The fore legs were brought up so as to brush but not touch the snout. (D44, No. 32) Stimulation led to movement of the upper lip with accompanying fore- and hind-leg movements. In one instance after some delay the fore leg was brought up with spreading toes brushing by the spot stimulated but not touching it. (D45, No. 33) Contraction of that half of the upper lip touched followed stimulation. The gum itself may not have been stimulated. (D46, No. 34) A sharp contraction of the upper lip as in a snarl, a pulling in of the tongue, and mandible elevation followed stimulation. Both fore legs were brought up to the side of the face. The hind legs were flexed to the belly with curving in of the rump resulting from flexion in the lumbar spine. (D47, No. 35) A localized contraction of the muscle of the lip followed stimulation. Violent movements of the fore limbs and the withdrawal of the head by the flexion of the neck and trunk muscles also took place. (D47, No. 36) A slight contraction of the lower lip and violent flexion of the left front limb followed stimulation. The mouth was then opened so as to pull back the upper lip. The tongue was then pressed forward rhythmically and the lower lip pulled down, bringing about a rhythmic, wide opening of the mouth. (D48, No. 37) Active tongue movements, lower-jaw movements, and pushing movements of the fore legs all followed stimulation. (D49, No. 38) Local contraction of lip muscles followed by rhythmic biting movements of the lower jaw occurred on stimulation. (D50, No. 40) Local response of the lip and a slight head and fore-limb movement followed stimulation. (D51, No. 41) The lips contracted in such a way as to bare the teeth after stimulation. (D52, No. 42) A pulling up of the lip on the side stimulated and a quick extension of the neck and consequent toss of the head followed stimulation. A slight homolateral fore-leg movement also occurred. (D53, No. 43) Contraction of the upper lip followed stimulation. (D54, No. 44) A local contraction of the segment with possible corrugation of the vibrissae pad followed stimulation. (D55, No. 46) Localized lip movements, an in-and-out movement of the tongue, a shake of the head, and a raising of the homolateral fore limb all followed touch stimulation. (D57, No. 48) Local movements of lips and tongue similar to those that would have been made in sucking and a single bite of the mandible were observed. (D58,

No. 49) N. R. (D59, No. 50) Twitches of both upper and lower lips and twitching of the fore legs occurred. (D60, No. 51) Pressure led to no response, but a needle stimulus led to the raising of both fore paws toward the mouth, all the while executing paddling movements. (D61, No. 53) There occurred a separate twitch of each half of the upper lip, tongue movements, and jaw movement. (D62, No. 54) Flexion of the fore legs and adduction and flexion of the hind legs, followed by a strong extension of the limbs and a slight movement of the head, all were noted. No lip movement occurred. (D63, No. 55) The left paw was raised to the point stimulated when the right gum was touched. When the left gum was touched, the right paw was raised to the stimulated spot. These responses were most precise in their localization. (D65, No. 58) N. R. (D67, No. 60) Continued rhythmic contraction of upper and lower lips was seen.

56. *Lower Gum.* (Summary only). At 35 days (No. 19), the first time at which a response to this area could be elicited, the following protocol is given: "Stimulation led to a raising of the right fore limb and a lowering of the head, so that the lips brushed along the right paw for the total distance from the toe to the elbow. No localized mouth movements could be observed." In the next recorded response (No. 24) lower-jaw and tongue movements, however, were elicited. After this, these two sets of movements, that is, limb raising and mouth activity, were the characteristic response to stimulation in this area. Apparently incidentally, other limb, trunk, and head reflexes were also released.

57. *Tongue* (Summary only). This proved a difficult point to stimulate without touching other parts of the mouth, but of those cases in which it was accomplished between the 37th day (No. 23) and the 65th day (No. 58) 16 cases occurred in which the tongue was retracted into the mouth after stimulation. In one case at 41 days (No. 29) the tongue was rhythmically drawn in and out of the mouth five or six times. Beside these local movements, incidental trunk, limb, and head responses were occasionally, but only occasionally, elicited.

58. *Pads at Base Front Digits, the "Interdigital Pads."*[6] (D32,

[6]Typically stimulation of this area consisted in touching with the side of the stimulus bristle the interdigital pads or eminences.

No. 12) Stimulation of the right fore paw when the animal was belly side up in the salt solution was followed by slight extension, that is, "fanning," of the digits of both fore paws. (D32, No. 14) A light touch of the pads of the right fore limb led to an active pressing of the rump to the same side. (D32, No. 14, F4) Stimulation led first to a flexion of the limb and then a thrust up against the stimulus. (D34, No. 17) Stimulation led to a trunk flexion in a bilateral "C" which also involved fore-leg flexion. (D35, No. 19) Light stimulation of the pad of the right fore limb led to a slight extension of the digits with a subsequent flexion of the apical joints of the digits and with a flexion articulated at the carpus. Stronger pressure stimulation at the same spot, however, led *to an extensor thrust pressing against* the stimulus as if to push it away. Again, the slight touch of the pad of the fore paw led to a slight movement of each of the other legs, such that, had it been *more extended, the legs would apparently have been brought together* in a point. (D36, No. 22) Stimulation led to a slight flexion of the digits. (D36, No. 21) A flexion of the fore leg to the side with a concomitant flexion and then extension of the wrist followed stimulation. The digits of the paws were slightly "cupped," i.e., flexed. (D38, No. 24) Stimulation led to a complete flexing of the digits so that their tips almost touched the proximal pads. (D39, No. 26) Touching the interdigital pads of the right fore leg elicited a sudden flexion of both fore legs and adduction of the limbs to the sides of the face. (D39, No. 25) N. R. (D40, No. 28) Stimulation led to a rhythmic extension-flexion of the paw at the carpus with an extension and fanning of all toes. (D41, No. 29) A slight flexion of the digits and a pressing of the limb toward *the stimulus* followed *stimulation.* (D42, No. 30) A withdrawal movement of the paw from the stimulus followed its application. (D42, No. 30, F2) A flexion of the stimulated fore leg toward the chest and a strong extensor thrust of the left fore leg with a fanning of the toes followed stimulation. The fetus was held ventral side up to the observer while these observations were made. (D43, No. 31) Sharp pushes against the stimulus and then upward paddling movements of the paw stimulated followed the experimental touch. A fanning or spreading of the digits also followed stimulation. After this the digits were flexed. (D44, No. 32) To

light touch no response was apparent, but with continued stroking of all the interdigital pads a flexion of the paw at the carpus with a cupping or flexion of all toes occurred. This was repeated. With continued stronger stimulation a greater cupping-flexion of the digits resulted. (D45, No. 33) Flexion of the stimulated leg into the chest with a slight extension at the wrist followed stimulation. Then paddling movements of both fore legs and a crossing of both fore legs with flexion at the wrist of the digits into a "fist" occurred. (D46, No. 34) A light touch led to a cupping or flexion movement of the toes with flexion of the paw at the wrist. Later a rhythmic extension of the paw with fanning and extension of the toes followed by a slight flexion and relaxation of the digits occurred. When the stimulus hair was held continuously parallel to the paw, definite "grasping" movements of the toes around the hair, with a slight pushing movement of fore arm against it, also occurred. After this the limb flexed at the elbow. (D47, No. 35) A flexion of the toes as in clasping, together with a flexion of the whole limb, followed stimulation. On repeating the stimulus a flexion of the digits occurred. This observation was repeated three times with similar results. (D47, No. 36) Stimulation, when the animal was held dorsal side down, led to small but obvious extension and flexion of all four limbs. (D48, No. 37) A slow flexion or grasp with the digits around the stimulus followed soft stimulation. Stronger stimulation led first to a slight flexion and then to a fanning and withdrawal of the paw from the point stimulated. (D49, No. 38) A quick extension of digits followed by a flexion of the digits which formed them into a tight fist occurred as a result of single momentary stimulus. The paw was then flexed at the carpus, while the digits were held extended. Only the paw stimulated responded; the other limbs or paws or trunk made no response. (D50, No. 40) Marked flexion of all digits of the stimulated paw was followed by flexion of the whole leg and then quickly by a similar flexion of digits and limb of the leg on the other side. Both limbs were then brought together so that the two paws stimulated each other. (D51, No. 41) An outward extension or fanning of the digits followed stimulation. (D52, No. 42) To soft stimulation no response occurred. On continued brushing, the digits slightly extended and fanned. Then alternate digit movements with articulation just

above the interdigital pads, similar to those used in playing the piano, took place. Wrist flexion and extension occurred after this. (D53, No. 43) Stimulation led to a pulling away of the whole leg with articulation at the shoulder. When stimulation was repeated extension of the toes only occurred. (D54, No. 44) A light touch led to flexion of the fore leg. With continued stimulation there was a flutter of the fore leg and a contraction of the opposite leg. The stimulated leg was brought slightly over and crossed the fore leg stimulated. A second stimulus led to a precise extension of the wrist and a flexion of the toes. (D54, No. 45) N. R. (D55, No. 46) No indication of digit movement followed even prolonged stimulation, but the whole limb twitched away from the stimulus. (D56, No. 47) N. R. (D57, No. 48) The first stimulus led to a flexion of the limb with articulation at the elbow. In this response the toes were flexed so as to cup them toward the stimulus. On continued stimulation the limb was withdrawn and the toes extended so as markedly to fan them. (D58, No. 49) Flexion of the leg with slight flexion at the wrist and fluttering movement of the toes nearest the point stimulated followed stimulation. (D60, No. 51) The pressure stimulus elicited no response. The needle stimulus led to a quick, strong flexion of the limb at the elbow, an extension at the wrist, and fanning of the toes. (D61, No. 53) N. R. (D62, No. 54) A sharp flexion of the stimulated leg followed the experimental stimulus touch. (D63, No. 55) N. R. (D65, No. 58) A slight flexion of the heterolateral limb, but no movement of the homolateral limb, followed stimulation. (D67, No. 60) Stimulation led to a sharp flexion and adduction of the stimulated leg to the chest.

59, 60. *Fore-limb Base Left and Right* (Summary only). These areas are relatively insensitive, especially in the later fetal stage. Flexion of the stimulated limb and of the whole trunk was released at 34 days (No. 17). From this time on flexion, or occasionally extension, of the limb followed stimulation. In the period between the 50th and the 67th day almost no responses were elicited in this region, even in response to needle stimulation.

61. *Epigastric Stroke* (Summary only). The term "epigastric stroke" was taken from clinical neurology, in which it is described as follows by Core: "With some pointed object . . . the skin over

the lower costal margin is lightly stimulated from above downwards. In the event of the *epigastric* reflex being present there is a twitch of the abdominal wall to the side stimulated; this is represented by a jerk of the median line to this side"(21).

At 34 days the first response to this form of stimulation was recorded (No. 17). It was a flexion of the rump and of both hind legs. At 38 days (No. 24) the first local muscle contraction took place. After this, hind-limb movement and local muscle contractions were characteristic responses until the 56th day (No. 47), after which, save for one slight hind-limb twitch at the 62nd day (No. 54) no response whatsoever occurred. Seven responses during the active period also involved the fore limbs and once (at 44 days, No. 32) one hind leg touched the stimulated area.

62. *Umbilical Cord* (Summary only). In the twenty-eight fetuses in which this was tried no response resulted, save that twice, at 48 days (No. 37) and at 51 days (No. 41) the cord was so strongly stimulated that it pulled on the abdominal wall, thus stimulating it, and in both cases then eliciting a flexion and adduction of both hind legs to the stimulated area. These results may be taken as showing conclusively that the umbilical cord itself is not sensitive during the period covered by this study.

63. *Umbilical Reflex Stroke* (Summary only). Like the epigastric reflex considered above, the conditions for this form of stimulation were adapted from the human clinical procedure as far as possible. Core says of this reflex, "In the case of the *umbilical* reflex the skin of the flank is stimulated in the same way from without inwards toward the median line. If the reflex is present, the umbilicus twitches to the side stimulated" (21).

The response to this form of stimulation began at 38 days (No. 24). The quotation of a single protocol at 44 days (No. 32) gives a picture of the typical response. "Stimulation led to a pulling in of the muscles at the base of the umbilical cord, followed by a flexion and adduction of the homolateral hind leg, such that its paw came in contact with the belly near the point stimulated." This was varied somewhat from fetus to fetus. Sometimes the heterolateral hind paw was raised. After the 57th day this reflex could be elicited only twice, but in both of these cases (61 and 62 days, Nos. 53 and 54, respectively) the response was characteristic.

64. *External Genital Region* (Summary only). The difficulty

of stimulating this region without also stimulating the anal region (No. 28 above) became apparent as the experiment progressed. The protocols could almost be exchanged with those given for No. 28. The dominant pattern here, as there, was one of local muscle contraction, rump flexion, and protecting thigh adduction.

65. *Plantar Stimulation*[6] (D34, No. 17) N. R. (D35, No. 19) Stimulation of the pad of the right paw led to a flexion of the rump muscles such that the rump was pulled in between the thighs. Stimulation of the pad of the left foot led to an extension of that hind leg with major articulation at the knee, but also with some articulation at the ankle. (D36, No. 21) No response even to continued strong stimulation, although the animal was in general very active. (D36, No. 22) N. R. (D37, No. 23) Stimulation led to a flexion of the digits. (D38, No. 24) Stimulation led to a flexion of the digits. (Note: Stimulating back of toes led to an extension of the digits which was quite the opposite of the flexion noted above.) (D39, No. 25) Stimulation led to a flexion of the leg with articulation at the heel and the knee. The stimulated limb was thus adducted to the belly. This was repeated with similar results. (D40, No. 28) A slight flexion of the limb first moved the paw away from the stimulus; an upward extension at the wrist accompanied by a fanning of the toes then occurred. This was followed by a flexion at the wrist and a slight flexion of the digits. (D41, No. 29) A pronounced fanning of the toes and an extensor thrust first followed stimulation. Pressing the stimulus harder against the interdigital pads then led to a definite flexion or cupping movement of the whole paw. It is remarkable, however, that light local punctiform stimulation of one pad at a time led to *fanning* or extension of the toes and especially of the toe above the stimulated pad. (D42, No. 30) Stimulation led to a slow flexion and adduction of the leg toward the belly. This flexion included an articulation at the heel. The limb then relaxed gradually back to normal position. (D43, No. 31) The paw flexed or curled slightly around the stimulus when it was applied. It then remained curled for a short interval of

[6]Typically stimulation of this area consisted in touching with the side of the stimulus bristle the interdigital pads or eminences. Sometimes the stimulus was brushed over the plantar surface more generally.

time after the stimulus was taken away. Articulation at the separate joints of the phalanges could be clearly seen in this flexion. (D44, No. 32) An extensor thrust of the leg, followed by a flexion or cupping of the toes, resulted from stimulation. (D45, No. 33) Stimulation led to a flexion of the stimulated leg such that the leg was adducted to the belly. The digits were flexed strongly. Touching one paw also led to a flexion of digits on the other. (D46, No. 34) At first a flexion of the limb followed stimulation. Then the digits were flexed and maintained in this posture for a time until this posture was, without further stimulation, overcome by an opposed extension, which came eventually to dominate completely so that a most vigorous fanning took place. (D47, No. 36) Stimulation led to a gentle flexion at the ankle. The three digits of the paw were also flexed. On repeated stimulation the magnitude of this response decreased. (D47, No. 35) A flexion of the toes around the stimulus object followed the application of the stimulus. Touching the interdigital pad at the base of all three toes of the paw led to an extensive movement of the two outer toes and a less pronounced movement of the middle one. Further stimulation led to an extension of the limb, thereby pulling it away from the stimulus. (D48, No. 37) An extension of the digits of both paws followed stimulation of the pads of either paw. (D49, No. 39) Stimulation at the base of each digit led to a flexion of that digit. Touching the paw a little further toward the heel, that is, at the proximal pads, led to a flexion of all toes and indeed to a flexion of the whole limb. (D49, No. 39, F4) A slow flexion and adduction of the whole limb toward the belly took place, after which followed a strong, vigorous extension of both hind legs and an extension of the rump. (D49, No. 38) Extension of the digits with articulation at the base of the digits and at the ankle took place. At the same time the homolateral fore limb was so flexed and adducted that it almost touched the point on the hind paw that had been stimulated. (D50, No. 40) Stimulation led to a flexion and cupping of the toes. When the pads at the base of the individual toes were stimulated there was a localized flexion. This observation was most marked in the case of the "little" toe. (D51, No. 41) A sharp flexion of the whole limb, a slight flexion of the digits of the stimulated paw, with individual twitching of digits, all followed a single touch stimulus.

(D52, No. 42) An extension of the toes followed light pressure stimulation. Stronger stimulation led to a flexion of the limb and a flexion and cupping of the toes. (D53, No. 43) Stimulation led to a strong quick extension and spreading of the toes. A second stimulus led to a similar response, including an extensor thrust of the whole limb. (D54, No. 44, F1) N. R. (D54, No. 44, F2) Stimulation led to a flexion of the leg and an extension of all toes. (D54, No. 45) N. R. (D55, No. 46) A pronounced flexion of all toes, followed by an extension of the hind leg, took place. General fore-limb, head, and trunk reactions also occurred. (D56, No. 47) Stimulation was followed by both hind legs being flexed and adducted to the body. (D57, No. 48) Stimulation was followed by a rhythmic flexion-extension movement of one toe. (D58, No. 49) A spreading and extension of the digits, then a flexion and bringing together of the digits, and finally a flexion at the ankle all followed stimulation. A light touch at the base of any one toe led to a flexion of that toe alone. The middle toe was most responsive. (D59, No. 50) A strong outward thrust of each hind leg followed stimulation of the interdigital pads of that limb. A very light touch, however, led to an extension of the toes. When the hair was thrust between the toes, however, the toes extended in a fan and at the same time the whole limb executed an extensor thrust. (D60, No. 51) No response occurred to pressure or normal needle thrust. Strong needle thrust led to an extensor thrust of the limb and a rhythmic flutter of the foot. (D61, No. 52) Stimulation was followed by a flexion of the toes and a slow flexion of the whole limb. (D62, No. 54) A sharp upward rhythmic extension of the hind leg involving extension at the ankle and at all of the digits followed stimulation. The same response was secured from the other leg. (D63, No. 55) A slight twitching backward of the foot with extension and fanning of all toes followed stimulation. A repetition of the stimulus led to a flexion of the limb which continued until the posture was such that further retraction was impossible. Then a violent thrust against the stimulus took place. This was observed four times. (D67, No. 60) No response to stimulation of any sort could be elicited.

66. *Blow for the Achilles Jerk* (Summary only). Although a number of responses are recorded to this form of stimulation, the

small size of the fetus and the difficulty of determining whether the response was one to mere cutaneous pressure has made it seem desirable not to include any report of these observations in the present paper.

67. *Leg (the Inner Surface of the Thigh)* (Summary only). This area proved to be relatively insensitive. At the 38th day (No. 24) occurred the first response, a flexure of the stimulated hind limb. This proved to be the characteristic response, but, save for occasional twitches, no response occurred to pressure after the 53rd day. A few extensor thrusts did occur, as well as incidental trunk and fore-limb reflexes.

68. *Cremasteric Stroke* (Summary only). The term cremasteric stroke is here used by analogy only. The stimulus used in human neurological examinations to elicit the cremasteric reflex is roughly similar in locus to the stroke indicated in Figure 1. This form of stimulation in the fetus elicited comparatively few responses. Of those that did appear, hind-leg flexion seemed most characteristic. In fetuses from the 51st to the 67th day in but two cases were responses recorded.

69. *Blow for Jaw Jerk* (Summary only). Because of the size of the fetus and because of the liquid environment it proved impossible, for the most part, properly to administer this form of stimulation. In one case quick jaw opening did result, but this was possibly the result of mere cutaneous stimulation.

70. *Taste Stimuli.* Several tentative experiments have been made in the study of taste stimuli by observing differences in the sucking reactions called out in late fetuses by different liquids, but these experiments are not well enough advanced for report at this time.

71. *Probe in Mouth.* (D34, No. 17) No mouth movements resulted from this form of stimulation, but both fore paws were brought up to touch the probe. The digits of both paws were extended and fanned. After removing the stimulus the paws were held at the point of previous stimulation. (D35, No. 20) Pushing a thin bristle probe in and out of the mouth led to several movements both of the jaw and the lip. The fore legs were also raised to the face. (D36, No. 21) The insertion of the bristle probe into the mouth led to a raising of one fore limb with flexed or cupped

digits. The paw then actually pushed against the stimulus object. After continued stimulation by moving the probe in and out of the mouth rhythmic movements of the lower jaw were initiated, which continued as long as the probe was gently moved in and out of the mouth. Tongue movements accompanied this activity. (D36, No. 22) Stimulation led to a wink of the unopened eye, mouth and lip movements, and a pushing of both fore legs against the face. (D37, No. 23) When the hair probe was placed in the mouth, rhythmic but very slow movements of the lower jaw began. (Note: In this animal self-stimulation of the inside of the mouth by the toes of the left fore paw was observed. This stimulation led to sucking movements, together with tongue movements.) (D38, No. 24) Stimulation led to rhythmic jaw movements. Soon after the insertion of the probe both fore paws were brought up as if to press away the stimulus. (D39, No. 25) The probe at first led to no jaw or tongue movements, but the lower jaw was depressed and then both fore paws were brought up to the mouth region. The mouth then began to open in a more or less rhythmic fashion, synchronized with a retraction of the muscle wall in the abdominal region. Then both fore legs were raised toward the head and the hind legs were extended away from the body. At the same time the neck was flexed, thus bowing it upon the chest. After a rest period the probe was again inserted in the mouth. This time tongue movements with the tongue slightly pursed and lip movements were also elicited. This whole pattern seemed to be very typical of sucking, but no indication of swallowing was present. (D39, No. 27) Stimulation led to upper-lip and lower-lip movements. The tongue was rhythmically protruded and retracted. Jaw movements also took place. (Note: In random behavior of the fetus, when the mouth came in contact with the placenta, mouth movements as described above were also observed.) (D40, No. 28) Upper- and lower-lip movements followed stimulation. Tongue and jaw movements also occurred rhythmically in such a way as to exert some pressure on the bristle probe. (D41, No. 29) Definite sucking movements of tongue, lips, and jaw, including hollowing of cheeks, followed stimulation. (D42, No. 30) Placing the bristle probe in the practically open mouth of the fetus was followed by a closing of the mouth and a pressing of the tongue around the hair. However, nothing that

could be called sucking movements followed. (D43, No. 31) Stimulation led to lower-jaw movement. The tongue was pursed over the object. Some rhythmic activity of the mouth was observed, but not such as to involve the throat. (D44, No. 32) Definite movements of opening and closing of the mouth, of the tongue, of the lip and jaw, strongly suggesting sucking or biting, followed stimulation. (D46, No. 34) Jaw, tongue, and lip movements with fore legs brought up to either side of the mouth followed stimulation. The fore legs then executed alternate paddling movements with fanning of the toes. (D48, No. 37) Upper- and lower-lip and tongue movement, all coordinated together in a premature form of sucking, followed stimulation. Definite pulling in of the rod in the mouth could be detected. (D49, No. 39) As soon as the probe was placed in the mouth it was pressed out by the tongue. At the same time both fore legs were brought up to the mouth region and brushed by it. A wooden probe ⅛" in diameter was now substituted for the smaller probe, with the result that prolonged sucking movements of the lip and tongue took place, but so far as could be observed externally no swallowing movements occurred. (D49, No. 38) Rhythmic jaw and tongue movements occurred in response to the probe that had been inserted in the mouth, the jaw was pulled back and out in such a way as to bring both upper and lower jaw to bear on probe. The lower lip then curled down, the upper lip was raised, but no rhythmic movements took place. (D51, No. 41) Tongue, upper- and lower-lip, and jaw movements all followed stimulation. (D52, No. 42) When the probe was inserted the tongue was protruded and the lips brought up and then retracted away from the teeth. The head moved in and out slightly and both fore limbs were brought up toward point stimulated. In one instance the tongue was thrust forth from the mouth. (D53, No. 43) Stimulation led to movements of lips and jaw though not in a typical "sucking rhythm." (D54, No. 44) No sucking movements followed insertion of the probe but rather a slight upward and downward rhythmic movement of the fore paws and a twitch of the trunk. Lip and tongue movements then took place, but not such that they could be called sucking. (D54, No. 45) Stimulation led to no true sucking movements, but lip movements did occur. The fore legs were brought up to the mouth and the head jerked back-

ward. (D55, No. 46) Stimulation led to rhythmic lower-lip movements. The tongue curved around the probe, was rhythmically protruded and retracted. At the same time the lower lip was raised and lowered rhythmically. (D56, No. 47) No lip or mouth movements followed stimulation. The fore legs, however, were brought up toward the mouth and the head was jerked away. (D57, No. 48) Stimulation led to localized lip and tongue movements. The neck was also extended so as to draw the head back and the two fore legs were flexed over the chest. (D58, No. 49) Stimulation led to tongue movement, lip movement, and jaw movement, all in a rhythmic sequence which may be characterized as sucking. (D59, No. 50) Lip, mouth, jaw, and tongue movements, all rhythmically repeated as in sucking, followed stimulation. (D60, No. 51) N. R. (D61, No. 53) No complex sucking movements could be elicited, but the lips did move. (D62, No. 54) Spontaneous mouth, jaw, lip, tongue, and throat movements were seen. Stimulation led to sucking movements which included clinging with lips, pressing with tongue, and even a slight pull on the probe. No indication of swallowing, however, could be observed. (D63, No. 55) Stimulation led to a biting down on the probe, quivering lip movements, and pronounced sucking and biting movements. At the same time both fore paws were raised and then extended so as to push away the stimulus. (D67, No. 60) Stimulation led to definite sucking movements, including lower lip, jaw, and upper lip. Pursing and pressing of the tongue in and out also occurred, as did neck movements, which seemed to be those of swallowing.

72. *Magnus Head Turning.* (Summary only). Under this somewhat inaccurate title are included the results of stimulation resulting from bending or twisting the head in relation to the trunk. In each case when the condition of the fetus made it possible the head was rotated first to the right and then to the left. After this the head was laterally flexed to right and left; it was then flexed dorsally and ventrally. The results of this form of stimulation in small fetuses are difficult to evaluate. At 37 days (No. 23), for example, rotation to the right seemed to lead to a slight extension of the right fore leg; rotating to the left, of the left leg. But these leg movements may possibly have been mere mechanical tensions brought about by the forced movement of skin and muscles. To all

of the 17 positive responses recorded, the same objection may be raised. According to the records secured, however, the very response given at 37 days was characteristic of later periods.

73. *Photic Stimulation of the Eye.* (In the cases marked "operation" the eye was opened by lifting the lids with cataract forceps and cutting the lids open. The photic stimulus consisted in turning on a 500-watt bulb at about 50 cm. from the eye.) (D39, No. 25) After the operation it was found that the eye was so deeply pigmented that the observation could not be made with assurance (D40, No. 28) After operative opening of the eye, eyeball movements could be observed, but in no direct relation to the photic stimulus. No iris responses were observed. (D41, No. 29) After operation and the flashing on and off of the light, definite but slight iris reflexes were observed. The pigmentation of this eye was favorable for observation. (D43, No. 31) After operation the response could not be surely determined as the eye was unpigmented. (D45, No. 33) The pigmentation of this eye was favorable for observation and after operation iris contraction to light stimulation could be seen. Marked eyeball movements also followed stimulation. (D46, No. 34) After operation marked eyeball movements were noted, but not in relation to light. Three observers all agreed that a definite iris reflex followed the photic stimulation, however. (D48, No. 37) After the operation marked eyeball movements were obvious, but light did not lead to an observable iris reflex. Eye winking was observed, but this was not surely related to the photic stimulus. (D49, No. 38) After operation violent eyeball movements were noted, but because of unfavorable pigmentation the iris could not be clearly seen and so the light reflex could not be studied. (D51, No. 41) After operation it was found that pigmentation made observation difficult, but iris response to photic stimulation did apparently take place. (D52, No. 42) After operation the pigmentation of the exposed eye was seen to be too great to make the iris visible. (D53, No. 43) After operation the reflex was impossible to observe because of unfavorable pigmentation of the eye and iris. (D55, No. 46) High pigmentation made observation of the iris impossible. (D57, No. 48) Pigmentation of this naturally open eye was such that no response could be seen. (D58, No. 49) Color of eye was such that observation was impossible with our technique.

(D59, No. 50) No iris response could be observed to photic stimulus. (D60, No. 51) Photic stimulus led to a pronounced slow contraction of the iris. (D61, No. 53) The iris could not be observed because of pigmentation, but the lids closed when the light was on and then slowly opened when the light was out. (D62, No. 54) The iris and eye were so darkly pigmented that the observation could *not* be made. (D63, No. 55) No response could be observed under the surface of the bath, but later in air the pupil contracted under the strong photic stimulus and dilated in weaker light. This response was repeated many times.

74. *Blow for the Triceps Jerk.* Stimulation proved not to be possible and no records were taken.

75. *Auditory Stimuli* (Summary only). In no case did the loud lung-blown whistle which was used as a stimulus lead to response while the fetus was still under water. In one case, in a 63-day fetus (No. 55) which was removed from the water, in 10 seconds a pinna reflex was noted. This was not tried in many cases and probably would have been effective in some younger fetuses. This response will be studied in more detail at a later time.

76. *Smell Stimuli.* In this study this form of stimulation was not systematically employed. Preliminary studies suggest, however, that it could be used differentially in those fetuses in which air breathing could be initiated.

77. *Needle Stimulation of Vibrissae Pad "B".* (In general, the needle, fastened to the end of a probe, was applied by a light touch sufficiently strong to be reported as a prick by a human subject. In certain cases it was, as noted, applied more strongly.) (D34, No. 17) Stimulation led to a sharp upward movement of both fore legs and brushing of the side of the face on both sides. (D35, No. 20) Slight movement of fore legs in general direction of stimulation followed application of the needle. (D36, No. 21) Stimulation led to quick neck bending so as to move the head away from the needle stimulus. This was followed by a unilateral bending of the trunk and an extension of the fore limb. The paw was brought very near the point stimulated. (D36, No. 22) A slight upward movement of the homolateral fore leg followed stimulation. (D38, No. 24) Stimulation led to an eye wink of the still unopened eye and was then followed by the paw's being raised toward the point

stimulated. The head was also raised as a result of neck extension. (D40, No. 28) N. R. (D41, No. 29) Writhing of the trunk muscles with flexion and extension of all four limbs followed stimulation. (D43, No. 31) The touch of the needle led to violent rotation of the head and opening and closing of the mouth. The homolateral fore leg was then brought up so that it brushed the stimulated spot. (D44, No. 32) Stimulation led to a sudden dorsal flexion of the neck and consequent moving away of the head, followed after several seconds by a slow precise raising of the homolateral fore leg to the exact spot stimulated. (D46, No. 34) A somewhat delayed response of the homolateral leg took place to stimulation, such that the paw actually pushed against the needle which was still in contact with the skin. After this, the homolateral fore leg was raised so that it rested between the eye and ear and then was brought slowly down the face until the paw was held over the snout where it remained "protecting" the recently stimulated spot for some time. (D48, No. 37) The pressure stimulus led to local twitch of the pad; so did the needle. (D49, No. 38) Stimulation led to local contraction of the muscles and a slight twitch of the fore limb and vigorous trot responses of the homolateral hind limb. The body was then flexed so as to bring the hind limb toward the point stimulated. The fore leg was then "curled" under the chest with flexion at all joints. Brushing the vibrissae now growing from the vibrissae pad in such a way as not to touch the skin elicited a more violent response than actually strongly pressing the skin at the base of the vibrissae with the needle. (D51, No. 41) After stimulation both fore legs were brought up and brushed past the snout. (D52, No. 42) The neck extended after stimulation and the posture was maintained for some time. Then all four limbs were extended away from the body. Several eye winks of the still unopened eye took place also. (D53, No. 43) Slight extension of the homolateral fore leg followed stimulation. A repetition of the stimulus led to a mere twitch of the homolateral limb. (D54, No. 44) The stimulus was applied with increasing pressure till the skin was pierced, but no response other than a twitch of the muscles underlying the pad occurred. The same response was given to a mere touch of the vibrissae themselves. (D54, No. 45) Stimulation led to a quick upward thrust of the left fore leg and then two or three

alternate paddling thrusts. (D55, No. 46) Local contraction of the pad followed stimulation. Each separate hair seemed to be independently erected. Then a pronounced writhing of the trunk muscles took place, and the homolateral and heterolateral fore paws were brought up to the head. The hillocks at the base of each hair became white as the vibrissae were erected. (D56, No. 47) Stimulation led to a strong rotation of the head, an upward thrust of the homolateral fore leg, and a slight raising of the homolateral hind leg. (D57, No. 48) Stimulation led to a twitch of the muscles of the neck. (D58, No. 49) No response to stimulation, no matter how strong the stimulus. (D59, No. 50) Stimulation was followed by the homolateral fore leg's being brought up sharply so as to touch the spot stimulated. The toes of the active limb were sharply flexed. (D61, No. 53) Stimulation led to a flexion of the neck, thus slowly bowing the head of the fetus and removing it from the region where it had been stimulated. (D62, No. 54) The moment the pad was pricked the head turned sharply away from the stimulus. This was repeated again and again. (D63, No. 55) Stimulation led to a slow and maintained eye closing, slight head bowing, with short movements of the homolateral paw toward the point stimulated.

78. *Needle Stimulation of Brow* (Summary only). Stimulation of this sort led to no responses until the 36th day and at that day only in one fetus for one of the two litters studied (No. 22). The response was a slow rotation of the head and a twitch of both fore legs. At later periods the responses released were mainly the special head reflexes. In no case were the responses noted as especially vigorous or strong.

79. *Needle Stimulation of Side-back* "C" (Summary only). Lateral flexures and dorsal and ventral extension and flexion of the trunk, including neck and rump, seemed to be the most characteristic responses released by this sort of stimulation. Rhythmic movements of the limbs were also elicited in many cases. This response was indeed the first to be observed as a result of this form of stimulation at 34 days (No. 17). At the 62nd and 63rd days (Nos. 54, 55) scratch reflexes were released. No indication of violent response or of the behavioral signs of pain were noted.

80. *Needle Stimulation of the Rump* (Summary only). No

response to this form of stimulation was elicited till the 38th day (No. 24). At that time the protocol reads, "Stimulation released paddling movements of all four paws and a flexion of the rump region such that the rump was pulled in between the hind legs." This is very typical of all the later responses to the stimulation of this area by the needle. None of these responses was especially vigorous.

81. *Needle Stimulation of Abdomen Side "A"* (Summary only). This area was relatively insensitive to stimulation. Adduction to the belly of the homolateral hind limb, activity of the other limbs, and local muscle response were characteristic reactions. The limb response pattern appeared at 36 days (No. 22), the local muscle response at 43 days (No. 31).

82. *Stimulation of the Front of the Nose by a Single-break Shock* (Abbreviated protocol). (Note: In order to use electrical stimulation for areas 82 to 93 inclusive small platinum-wire electrodes were applied to the area to be stimulated, which was for the purpose held above the surface of the bath. However, the surface of the animal was always still moist. In the case of the single-break shocks and of the faradic current the inductorium was adjusted so that the shocks were just perceptible to the human subject. Some further adjustment was at times made after the electrodes had been applied to the fetus. In all cases the response from *touching* the fetus by the electrodes was allowed to dissipate itself before the electrical stimulus was given. For early responses to this sort of stimulation see first sections of protocols.) (D34, No. 17) A slight upward movement of the head, accompanied by a brushing movement of the fore leg, followed stimulation. (D38, No. 24) After stimulation the mandible was elevated and all four legs twitched. (D40, No. 28) An upward thrust of both fore legs followed stimulation. (D41, No. 29) A slight contraction apparently involving most of the muscles of the trunk followed stimulation. (D43, No. 31) Stimulation led to an upward movement of fore legs toward but not touching the snout. (D45, No. 33) A wink of the still unopened eyes and a sharp flexion of the homolateral hind leg followed the application of the stimulus. (D46, No. 34) A wink of the still unopened eye followed stimulation. The homolateral fore leg was then flexed and the limb brought up to the side

of the face. The paw flexed during this movement. (D49, No. 38) Stimulation led to the head's being slightly thrust forward. The nostril then dilated and contracted and the fore paws were twitched. (D51, No. 41) A twitch of the musculature around the electrodes and a wink of the still unopened eye followed stimulation. (D52, No. 42) The trunk and limbs executed a quick writhing movement which especially involved the homolateral leg, which was adducted to the nose but not so as to touch it. (D53, No. 43) A rotation of the head away from the stimulus and an extension of both fore legs away from the body resulted from the shock stimulus. (D54, No. 44) A localized corrugation of the vibrissae pad followed stimulation. (D54, No. 45) A violent extension of the neck muscles and backward toss of the head followed stimulation. (D55, No. 46) Stimulation led to a jerk of the head. (D56, No. 47) An upward thrust of the heterolateral fore leg and a turning away of the head followed stimulation. (D57, No. 48) Contraction of the muscles about the eye and an ear twitch both followed the stimulus. (D59, No. 50) A violent twitch and twist of head and a violent outward thrust of the legs occurred after stimulation. (D60, No. 51) A twitch of the head followed stimulation. (D61, No. 53) Responses involving trunk and limbs in a typical "struggle" reaction followed stimulation. (D62, No. 54) A sharp upward thrust of the fore leg followed stimulation. (D63, No. 55) Stimulation led to a dilation of the nostril, together with a slight flexion of the fore paw.

83. *Stimulation of Area "B" of the Pinna by a Single-break Shock* (Summary only). Because of the size of the electrodes no stimulation of this area was tried before the 37th day (No. 23). This animal gave a local contraction of the pinna, which may well have been a direct muscle response. In later fetuses the pinna reflex was still the characteristic form of response to this sort of stimulation, but general head movements, the other head reflexes, and indeed limb and trunk movements were also occasionally elicited.

84. *Single-break Shock Stimulation of Side of Back "B"* (Summary only). Local muscle response, flexion or extension of all limbs, and general trunk movement followed this form of stimulation. In a few instances the homolateral hind leg was adducted to the stimulated spot.

85. *Single-break Shock Stimulation of Abdomen Side "B"* (Summary only). Following application of the stimulus there was a local response of the muscles near the electrodes and an assortment of trunk and limb movements, which did not seem to follow any particular pattern. The first response was at 38 days in No. 24.

86. *Angle of Lip* (*Faradic*) (Abbreviated protocol). (Note: Stimulation given as described under No. 82 above.) (D34, No. 17) Stimulation led to violent paddling or trot movements with all four legs. (D40, No. 28) General trunk and limb responses occurred. (D41, No. 29) A localized brushing movement of the homolateral paw touching the point stimulated followed the brief stimulus period. (D43, No. 31) General trunk activity followed stimulation. (D44, No. 32) A downward extensor thrust of the homolateral fore leg followed stimulation. (D45, No. 33) After stimulation the tongue was retracted, the lower lip twitched, and the homolateral hind leg was brought up sharply to the belly. (D46, No. 34) Corrugation of the vibrissae pad, wink of the still closed eye, and a twitch of all four legs followed stimulation. (D49, No. 38) An elevation of the lip and flexion-extension sequences of movement in both fore legs resulted from stimulation. (D51, No. 41) Stimulation released general body and limb responses. (D52, No. 42) Pronounced responses of the trunk and all limbs were released by the current. (D53, No. 43) An extensor thrust of the hind and fore legs followed stimulation. The head then turned and the local musculature contracted. (D55, No. 46) Stimulation led to a quick jerk of the head and general mouth movements. The fore paw was then brought up to the spot stimulated and both hind limbs were extended and the trunk gave a twitch. (D56, No. 47) Stimulation led to mouth movements and a turning away of the head. After this both forelegs were brought up and extended, and the hind legs were extended away from the body. (D57, No. 48) Eye wink followed stimulation. (D60, No. 51) Extension of all four legs and of the neck followed stimulation. (D61, No. 53) General struggle response resulted from stimulation. (D62, No. 54) Stimulation led to a contraction of the local musculature. The homolateral fore leg was flexed to the chest and the head extended and thus pulled away. (D63, No. 55) Stimulation led to a head twitch, jaw movements, the raising of the fore paw, and a twitch of both hind legs.

87. *Faradic Stimulation of Side Face "B"* (Summary only). At 40 days (No. 28) wink of the still unopened eye, neck and fore leg flexion took place. This was the characteristic pattern of response released by stimulating this area in all later fetal stages.

88. *Faradic Stimulation of Area "A" of the Pinna* (Summary only). At 34 days (No. 17) general limb movements were released. At 40 days (No. 28) pinna reflex and eye wink of the still unopened lids occurred. From that time on the elicitation of the head reflexes and of fore-limb movement was characteristic.

89. *Faradic Stimulation of the Back of the Shoulder* (Summary only). Local contraction of the skin near the electrodes, trunk flexion, and movements (flexion or extension) of all four legs typically followed stimulation here after the 40th day.

90. *Faradic Stimulation of the Rump* (Summary only). At 40 days (No. 28) the response recorded to this form of stimulation is: "A local contraction of muscles, an inward curving of the rump, and a slight flexion and then extension of the hind legs." This was characteristic of the responses to stimulation in this region up to the last recorded observation in a fetus of 62 days. This pattern of behavior was similar to that elicited by stimulation at Nos. 26, 27, and 80.

91, 92. *Faradic Stimulation of Hind-foot Toe "C" and Front-foot Toe "A"* (Summary only). In both of these areas the typical response was either a flexion or extension of the toes and either a flexion or extension of the stimulated limb. Other limb and trunk movements occasionally occurred. There was no indication of especially strong responses.

93. *Faradic Stimulation of Tongue* (Summary only). The mechanical difficulties of applying electrodes in this area make the results secured by this sort of stimulation questionable. Trunk, limb, and local mouth movements were all elicited, however.

93. *Faradic Stimulation of Tongue* (Summary only). The mechanical difficulties of applying electrodes in this area make the results secured by this sort of stimulation questionable. Trunk, limb, and local mouth movements were all elicited, however.

94, 95. *Temperature Stimulation at the Side of the Face (Area "D") and the Side-back (Area "C")* (Summary only). In order to stimulate the fetus with temperature, the organism was first raised until the body area to be stimulated was just above the

surface of the bath. A few drops of water from the bath were then dropped upon the skin to test for pressure responses to the falling drops. The drops were allowed to fall a few mm. only. In the 34 experimental stimulations only three or four times did any response follow this form of stimulation and in these cases it was very slight. The temperature stimuli used were relatively intense. In a later study less marked temperatures will be studied. In the present experiment cold stimuli of 2° and 4° C. and warm stimuli of 70° and 75° C. were used for the most part, although response to temperature of 19° C. was recorded in one case. The responses released by the cold and warm stimuli were for the most part similar and involved many muscles of the trunk and limbs. There was a typical difference between the stimulation of the face and the side, in that the face stimulation did in many cases lead to the raising of the homolateral fore limb toward the place stimulated. Unmistakable responses to temperature stimuli were secured in typical fetuses from the 35th day to the 62nd day. It must be noted that in each stimulus recorded above pain may have been involved because of the extremes of temperature used.

96, 97, 98, 99, 100. *As Stimulated after High Cervical Section of the Spinal Cord* (Summary only). Without interfering with the placental blood supply, a quick incision was made in the neck in certain of the fetuses studied and then the whole spinal column and cord was severed by sharp cut of a small pair of surgical scissors. The fetus ordinarily died in a short time after this very crude operation, presumably because of loss of blood. The many serious faults with the technique described are obvious. The results given below are therefore presented as suggestive rather than as satisfactory experimental findings. The cord section was usually in the mid-cervical region. In the summary of findings presented here no effort will be made to follow the areas of the chart, but the behavior of each fetus studied will be briefly summarized.

(D37, No. 23) At the moment of cutting the spinal cord a definite, pronounced, simultaneous flexion of both hind legs occurred. The lids of the still unopened eye responded by winking when directly stimulated. For approximately 30 seconds after transection of the cord no reflexes of those tried could be elicited, save the orbicularis just noted. After this, limb responsiveness returned. Stim-

ulation of the pads at the base of the paws (fore and rear) led to extensor thrusts of greater magnitude, it seems, than in the intact animal. When the fore legs were mechanically bent they returned the pressure in an exaggerated form. (D40, No. 28) After the cut there seemed to be increased tonus of the fore limbs. Reflexes were maintained in the hind legs. Eyelid and pinna reflexes were also maintained. No paddling or general limb movements, however, occurred. (D41, No. 29) After a very short interval following section both fore legs became rigid. This was not true of the hind legs. Reflexes could not be elicited. To passive flexion of the fore and hind limb, however, the position to which the limb had been bent was maintained. (D43, No. 31) After transection no movement of the pinna could be elicited, but the unopened eye did wink. Stimulation of the head region had no effect on the trunk or limbs, but stimulating the fore limb did lead to responses in the hind limb and stimulating the hind limbs led to a flexion in the fore limbs. Trot reflexes of long duration were then initiated in the fore limbs and the hind limbs stretched out from the body holding long-maintained extensor tonus. (D44, No. 32) After transection there was an exaggerated extensor thrust of both fore legs, when touched with the stimulus hair on the bottom of the interdigital pads. There was also a quick active extension and fanning of the toes. Virtually the same condition prevailed in the hind legs. Movements seemed to be much more intense, vigorous and uncontrolled than they had been in this fetus before transection. (D45, No. 33) After transection stimulating the angle of the lips led to a wink of the still unopened eye. To a stimulation of the hind limb an extensor thrust occurred, which was also followed by an extensor thrust of the heterolateral fore limb. The hind limbs were then both extended and maintained with a marked rigid tonus. Stimulating the elbow of one fore limb led to a flexion of the heterolateral hind limb. (D46, No. 34) After transection an extreme rigidity of both hind legs took place. The right fore leg was then extended and brought down to and touching the hind leg. The left fore leg was then flexed and maintained flexed. After this rigidity passed from the hind legs. Stimulation of the unopened lids led to a "wink" although there was no response to stimulation of the concha. Very slight touch stimuli led to responses when the hair was applied to the hind legs, fore

legs, or abdomen. (D48, No. 37) After transection stimulation of the angle of the lip led to movement of the facial musculature and of the lips. Stimulating the hip led to a lateral flexion of the spine, and a touch of the interdigital pad of the fore paw led to a violent cupping and pushing movement with toe extension and shaking of the head. Strong hyper-activity in general characterized the animal. Passive flexion of the fore limb led to general "struggle" responses. (D49, No. 38) Immediately after transection trotting movements of all four legs began. These movements stopped in the hind legs first, but continued in the fore legs for some seconds. Exaggerated extension of the toes of the hind leg was elicited by stimulating the interdigital pads, and exaggerated flexion or cupping in the digits of the fore paws occurred to similar stimulation. In both of these cases there was no doubt that the response was exaggerated in comparison with any response elicited in this fetus previously. (D51, No. 41) After transection the hind legs were markedly extended and the fore legs executed paddling movements. In the rhythm of the fore-leg movements there was something of the repetition characterized as the "trot." Lower-jaw, mouth, tongue, lip movement, and definitely sucking responses followed stimulation of the corner of the lip. Stimulation by pressure on the wrist led to an extension of the toes of the stimulated fore limb. Stimulating the interdigital pad of the fore foot led to wrist extension and the fanning of the toes. Stimulating the same pad on the hind paw led to an extensor thrust and a rhythmic and alternate cupping and fanning of the toes. Stimulating the concha led to sucking movements, but no responses of limb or trunk. In general, a response such as a flexion of the toes was more pronounced than in the case of the normal fetus. (D52, No. 42) After transection pronounced rigidity of the hind legs, rapid paddling movements of the fore legs, which then passed over into a somewhat maintained rigidity, followed stimulation. Slight rhythmic action in hind legs was also noted. A touch of the interdigital pads of the hind paw led to a violent extension of both hind limbs with a marked flexion of all toes. (D53, No. 43) After transection stimulation led to continued paddling movements of both fore legs and an extension of the trunk below the cut. The fore-leg movements continued for a few seconds, then stopped. Touching the corner of the mouth led to a

contraction of the musculature around the nostril. An eye wink of the still unopened lids with a maintained contraction also occurred. Touch on the hip led to no response, not even to a deep pressure stimulus. A touch of the interdigital pad of the hind paw led to a slight extension of one digit. (D55, No. 46) Immediately after the cut a pronounced extension of the fore and hind legs took place which involved a rhythmic movement in the hind legs with pronounced extensor tonus maintained. A touch in the corner of the lip led to tongue movements. No response occurred to the touch of the hip, but to touching the pad at the base of the fore paw an extensor thrust took place which was followed by decreased tonus. No response resulted from stroking the abdomen. (D57, No. 48) After transection the paw reflexes were difficult to elicit, but did appear. In general, the movements were somewhat slower and with a more continued tonus than in the normal animal. (D59, No. 50) After transection no response to the corner of the mouth took place. Touching the hip led to a contraction of the muscles around the anus. A touch of the interdigital pad of the plantar surface led to a violent withdrawal of the limb. (D60, No. 51) After transection paddling movements of the hind legs took place. The fore legs were not active, but were still sensitive to stimulation. Head and trunk responses occurred to appropriate stimulation. (D63, No. 55) Touch stimuli did not elicit response. However, pronounced responses did occur to strong stimulation by forceps and by passive-flexion movements of the limbs. The section was high in the cervical region and the eyelid reflexes seemed to have been abolished, as were the pinna reflexes.

101, 102. *Rotation and Post-rotation in Air-breathing Fetuses* (Summary only). Several fetuses of an age to be able to breathe when placental circulation was interrupted were tested on a turntable. This experimentation was not done systematically, but there was no doubt of compensatory head and trunk movements, both before and after rotation in fetuses of 65 days. These responses will be further studied.

103, 104. *Passive Flexion of Limbs.* Incidental notes in regard to response to this form of stimulation have been given above. It is not given in detail here because it was always complicated by pressure produced at the point at which the limb was grasped to produce the flexion.

III

CONCLUSIONS

A number of protocols and summaries have been presented above, in as complete form as feasible in the available space, in order to make clear to the reader the complex nature of the behavioral life of a series of fetal guinea-pigs of increasing gestation age when such behavior is released at each stage by the stimulation of definitely located receptor areas. Table 2 presents an abbreviated summary of the general locus of the first response and the date of such response in relation to the special receptor areas stimulated.

The actual picture of fetal behavior gained by reading the protocols is more important and more true than any summarizing statements which may be offered. The conclusions presented below should, therefore, be evaluated in the light of this fact.

1. *Heart beat was the only activity observable in the youngest fetuses studied in this investigation.* This finding is in accord with the reports of many previous investigators upon comparable mammalian fetuses (76, 104, 15). For reasons that need not be reviewed here heart beat is not considered to be "behavior" as that word is used in this paper. The fact, however, that heart beat may lead to passive movement of the head in the cat fetus has been noted (101).

2. *Before any behavior was observed the skeletal muscles of the still immobile fetus could be made to respond by direct electrical stimulation.* This finding is in accord with the previous observations of a num-

TABLE 2

Receptor area stimulated	Age of fetus at first response	General locus of first response
1. Angle of lip	32	Neck and fore limb
2. Vibrissae pad "A"	32	Neck and both fore limbs
3. Vibrissae pad "B"	32	Neck, trunk and both fore limbs
4. Nostril	32	Neck and fore limb
5. Front nose	32	Fore limb
6. Side face "A"	34	Caudal trunk and fore limb
7. Side face "B"	37	Fore limbs
8. Side face "C"	32	Neck, fore and hind limbs
9. Side face "D"	34	Trunk, neck, and fore and hind limbs
10. Eye (when open)	56	Eye wink
11. Lower eye lid	32	Fore leg
12. Upper eye lid	34	Neck and fore leg
13. Brow	40	Wink of unopened eye and pinna reflex
14. Crown	34	Neck
15. Concha	31	Fore limb and neck
16. Pinna "A"	32	Trunk
17. Pinna "B"	37	Pinna and palpebral reflex
18. Neck ventral	32	Fore and hind limb
19. Neck dorsal	32	Fore and hind limb
20. Side above shoulder	34	Fore limb and shoulder girdle
21. Back above shoulder	34	Trunk, neck, fore and hind legs
22. Side-back "A"	32	Hind limbs
23. Side-back "B"	35	Local subcutaneous muscle response
24. Side-back "C"	38	Rump, fore and hind limbs
25. Side-back "D"	38	Rump, fore and hind limbs
26. Rump	36	Trunk, fore and hind limbs
27. Hip	32	Trunk, fore and hind limbs
28. Anus area	35	Fore and hind limbs
29. Knee	34	Fore limb
30. Leg	35	Fore and hind limbs
31. Foot "A"	38	Digit flexion
32. Foot "B"	35	Fore leg
33. H. toe "A"	34	Hind limb
34. H. toe "B"	*35	Digit extension, fore and hind limbs
35. H. toe "C"	*40	Digit and ankle movement
36. Abdomen side "A"	34	Hind limbs
37. Abdomen side "B"	32	Trunk
38. Shoulder	34	Fore limb
39. Elbow	34	Fore limb and digits
40. Fore arm	35	Fore limbs
41. Wrist	34	Neck and limb
42. F. toe "A"	32	Fore limb and neck
43. F. toe "B"	*41	Wrist and limb

TABLE 2 (*Continued*)

Receptor area stimulated	Age of fetus at first response	General locus of first response
44. F. toe "C"	*40	Digits and limb
45. F. toe "D"	*41	Digits, wrist, elbow, shoulder
46. Dorsal mid-line snout	35	Fore and hind limbs
47. D. M. L. crown	38	Head, fore paws
48. D. M. L. neck	35	Neck, fore and hind limbs
49. D. M. L. shoulder	32	Trunk, neck, fore and hind limbs
50. D. M. L. back	38	Local subcutaneous muscle, fore and hind limbs, neck, jaw
51. D. M. L. hip	38	Neck, fore and hind limbs
52. D. M. L. rump	38	Rump
53. Ventral mid-line snout	36	Fore limbs
54. Lip groove	35	Neck, fore limbs
55. Upper gum	32	Fore limbs
56. Lower gum	35	Neck, fore limbs
57. Tongue	36	Tongue movement
58. Pad at base F. digit	32	Fore limb
59. Arm base right	34	Fore limb
60. Arm base left	34	Trunk and fore limb
61. Epigastric stroke	34	Rump and hind limbs
62. Umbilical cord	—	No direct response at any time
63. Umbilical reflex stroke	38	Rump and hind limbs
64. External genital region	36	Subcutaneous muscles and hind limbs
65. Plantar stimulation	35	Rump, hind limbs
66. Blow for Achilles jerk	—	Difficulty of stimulation makes response in doubt
67. Leg	38	Hind limbs
68. Cremasteric stroke	38	Hind limbs
69. Blow for jaw jerk	—	Difficulty of stimulation makes response in doubt
70. Taste stimuli as noted	—	Difficulty of stimulation makes response in doubt
71. Probe in mouth	34	Fore limbs
72. Magnus head turning		(see protocols)
73. Light stimulus (Eye open naturally or by operation)	41	Iris response
74. Blow for triceps jerk	—	Difficulty of stimulation makes results in doubt
75. Auditory stimuli as noted	63	Pinna response
76. Smell stimuli as noted	—	Difficulty of stimulation makes response in doubt
77. Vibrissae pad "B" (Needle)	34	Fore limbs
78. Brow (Needle)	36	Neck and fore limb
79. Side-back "B" (Needle)	34	Fore and hind limbs
80. Rump (Needle)	38	Rump, fore and hind limbs

TABLE 2 (*Continued*)

Receptor area stimulated	Age of fetus at first response	General locus of first response
81. Abdomen side "A" (Needle)	36	Fore and hind limbs
82. Front nose (Single-break shock)	34	Neck and fore limbs
83. Pinna "B" (Single-break shock)	37	Pinna response
84. Side-back "D" Single-break shock)	34	Local muscle contraction, fore and hind legs, trunk
85. Abdomen side "B" (Single break shock)	38	Trunk, fore and hind limbs
86. Angle of lip (Faradic)	34	Fore and hind limbs
87. Side face "B" (Faradic)	40	Palpebral reflex, neck, fore limbs
88. Pinna "A" (Faradic)	34	Fore and hind limbs
89. Back above shoulder (Faradic)	40	Trunk, fore and hind limbs
90. Rump (Faradic)	40	Trunk, fore and hind limbs
91. H. toe "C" (Faradic)	40	Hind limbs
92. F. toe "A" (Faradic)	40	Fore limb and digits
93. Tongue (Faradic)	—	Difficulty of stimulation makes results in doubt
94. Side-face "D" (Warm and cool)	35	Trunk, fore and hind limbs
95. Side-back "C" (Warm and cool)	35	Trunk, fore and hind limbs
95. Angle of lip (Pressure, etc., after high cervical section)		(See special protocols)
97. Hip (Pressure, etc., after high cervical section)		(See special protocols)
98. Pad at base F. digits (Pressure, etc., after high servical section)		(See special protocols)
99. Epigastric stroke (Pressure, etc., after high cervical section)		(See special protocols)
100. Plantar stimulation (Pressure, etc., after high cervical section)		(See special protocols)

* In these cases the day given is the first day on which the area had been stimulated.

ber of investigations on comparable fetal material (43, 103, 65).

3. *In a 28-day fetus behavior involving skeletal muscle response without electrical stimulation was ob-*

served for the first time. This is in virtual agreement with the earlier work of Preyer and Yanase, who set the onset of motility in the fetus of the guinea-pig at about 4 weeks (76, 104).

4. *The first observed responses occurred "spontaneously," that is, as a result of unknown causes.* These responses were in fetuses in which muscular reactions had not as yet been released by experimentally applied exteroceptive stimulation. In other words, Preyer's generalization, *"die Sensibilität tritt regelmässig später auf, als die Motilität"* (76), was verified in the onset of behavior. This finding is in harmony with the observation on other fetuses made by previous investigators, although it is possibly not true for all fetuses (77, 93, 100). It is conceivable that the mechanical changes incident upon opening of the uterus in our preparation lead to proprioceptive (deep pressure?) stimulation of a sort not provided by our exteroceptive pressure stimulation. If this is the case, the "unknown cause" of the first spontaneous movements is proprioceptive stimulation and its definite "reflex" form is more easily understood. Windle, Orr, and Minear hold that the proprioceptors may be the first functional receptors in the cat fetus (103). Later "spontaneous" movements may result from interoceptive or other general activation of previously established response systems and thus present a more typical picture of what has sometimes been called spontaneous "mass behavior."

5. *The first "spontaneous" movement observed was a lateral flexion of the neck and a synchronous and possibly independent movement of the fore limbs.* This

finding is not absolutely conclusive in regard to a vexed problem of the present time concerning the origin of limb movement. Windle and his collaborators, as suggested above, hold that the first movement observed in the cat is of the nature of a specific reflex (101, 103). Coghill, Angulo y González, and others (18, 4), on the other hand, working with other fetuses, hold that the first movement is a generalized trunk response, later involving passive fore-limb movement and then independent limb movement. Windle and Orr find that the first spontaneous movements of chick embryos are ventral trunk flexions, but the first responses to exogenous stimuli in this organism, which is observed some days later, is a wing reflex (67, 102). The author has further experiments in progress which will assist in settling this point so far as the guinea-pig fetus is concerned.

6. *The first sensory area from which behavior was released in the present study was Region 15, the concha of the ear.* This occurred in a fetus of 31-gestation-days' age. At this time no other cutaneous area released responses, when stimulation was applied as it was in this study. Windle reports that in the cat fetus the "first reflexogenous zone" included ". . . . the nose, ears, and in general, most of the head . . . ." (100).

7. *In general, responses at every stage are a function of what were called in the introduction of this paper (a) modes of stimulation and (b) variable conditions of the organism.*

8. *The first stimulus-released response noted in the study may be characterized, at its first appearance, as*

a pattern of behavior, which involves a relationship between neck flexion and fore-limb movement. Moreover, this "pattern of behavior" proved to be characteristic of the "pattern" which was typically released by stimulating the concha during the whole period of fetal life. An observation similar to this seems to have been made by other investigators, but little emphasis has been previously placed upon its "patterned" character. It is quite possible, however, that neurological study will demonstrate that the head-movement and limb-movement aspects of the response are to some extent independent mechanisms at this age.

9. *Many of the points indicated on the Chart of Receptor Areas, when stimulated, release behavior from the first which in spite of very great variability could always be considered as a special "pattern of behavior."* A consideration of the protocols given above demonstrates this fact. The simultaneous presence of secondarily released behavior, as explained below, sometimes tends to "mask" this pattern. The term "pattern of behavior" as used here and in other parts of this study is not intended to imply in a sequence of behavior any mystical sort of "totality," "inner organization," or the like, but merely to characterize various relatively complex series of responses which are recognizably similar in some respect. Patterns are characterized as similar, that is, as partially identical, either because the elementary responses composing them are identical in absolute magnitude of movement and in absolute time relationships, or are recognizably proportional, at least in part, to each

other in the relationships existing between these times or magnitudes. In many ways the logic behind the construction of the concept of the "behavior pattern," when that term is used in a non-mystical sense, is comparable to the logic used by Sherrington in the development of his conception of the "type reflex" (81).

10. *Each behavior pattern released by the stimulation of a particular area may be said to undergo a series of changes during fetal life.* The statement just given must be recognized as a metaphorical and qualitative expression intended to represent the fact that while certain relationships called "a pattern" are maintained, at least relatively, other relationships are absolutely altered. It should be noted that the author does not consider the "pattern" to be "a thing," but merely as a linguistic descriptive device forced upon one who describes behavior in words, by the nature of language itself. The changes in patterns may thus be called alterations of *motor diffusion to specific receptor stimulation* at various ages. To divide artificially this intricate continuum, however, which is in a sense specific for each receptor area, into a set of discrete stages intended to fit all areas is to do an essential injustice to the facts. For the purposes of exposition, nevertheless, it seems that the device of constructing arbitrarily "stages" of change in "behavior patterns," provided such stages be recognized as pure constructs, may not be without value. It should be noted also that quantitative study of particular behavior sequences at any "stage" would show great variability. Eventually quantified stimulus conditions and measured responses

must be recorded before completely valid scientific conclusions can be drawn in regard to these so-called stages. With these qualifications it may be said that the most typical course of development of behavior in relation to specific sensory stimulation seems to involve the following five stages:

Stage A. The first release of behavior, in which, weakly and incompletely, certain aspects of what will later form the typical "pattern" of the *gross responses* from that area are released. By "gross responses" is meant large limb-muscle and trunk movements. If the reader will refer to the full protocols given above, he will see a number of clear examples of this and the following stages as exemplified in the stimulation of specific receptor areas. In recently reported work on the cat embryo Windle, Orr and Minear (103) have described certain very early movements of this sort as definite "reflexes." To the present writer this seems entirely justified. This observation of Windle and his collaborators and the observations of the present investigation seem sharply to qualify, therefore, the view that all fetal development of specific responses is the "individuation" of previously existing "mass behavior." The desirability of the use of the term *Reflex* in this connection is merely a matter of convention. However, ultimately the present writer subscribes to the view that such responses may well be called reflexes. Skinner has offered the following clear proposition in regard to the relationship between general behavior and the reflex: "The essence of the description of behavior is held to be the determination of functional

laws describing the relationship between the forces acting upon, and the movement of, a given system. The reflex is, by definition, the precise instrument for this description" (83).

Stage B. The pattern of gross movement which appeared weakly and incompletely in *Stage A* is now stronger and in many cases more "adaptive"; that is, for example, in the case of *area 15,* instead of a slight bowing of the head and a slight elevation of the fore limb, both of these responses are now carried out so completely that the point on the ear which has been stimulated may be actually brushed by the paw.

Stage C. Without any necessary change in the gross movement pattern considered in *Stage A* and *Stage B* a fine movement pattern ordinarily involving limited and usually small muscle-groups is now often released by stimulation which was not in any sense present in the behavior initiated by identical stimulation in previous stages. From this time on, however, this response becomes a typical component of the characteristic behavior pattern released by stimulating many specific receptor areas.

Stage D. As a result of the structural alteration underlying changing thresholds of stimulation, or as a consequence of a modification in other underlying structures and functions, especially those concerned with central-nervous-system facilitation and inhibition, the pattern of responses released by a given sensory area may now frequently become less complete than formerly. For example, in this stage, may often appear only the eye-wink component of the whole pattern

which was previously elicited by the stimulation of the lids. The stages outlined as *A, B,* and *C* are thus typical in the history of such definite reflexes as the *palpebral,* the *pinna,* and many others, as observed and recorded in this study. It should be understood, however, that the possibility of calling out the larger pattern of which these definite reflexes were once a component has not been lost, but may be effected by appropriate conditions of the organism or the stimulus at any time. In general, this analysis may be seen to be in a measure in accord with Coghill's description of the development of a reflex as *individuated* out of a total pattern (20). It should be noted, however, that typically not until our *Stage C,* above, has this elementary component which is later to be "individuated out" of a more complex behavior pattern been characteristically apparent at all.

Stage E. Type I. In many areas, but by no means in all (and notably not in mucous-membrane areas), in the latter part of fetal life receptors previously effective in eliciting patterns of behavior become ineffective. But such responses as are released are now often fast and precisely localized. The unexpected phenomenon of the apparent unresponsiveness of the late fetus may not be unrelated to one or more of the following synchronous changes in the organism: (1) Mechanically, in the manner emphasized by Kuo in the chick (52), the fetus has now become so large that movement is difficult. In many instances in late fetal life the organism, on being delivered into the bath, requires a good deal of stimulation before it seems to "awaken."

This is in sharp contrast to the immediate activity apparent under similar conditions in younger fetuses. (2) The skin has become thicker and the hair coat more protective at this period. That this is not the only factor involved is attested, however, by the observation that needle stimulation as well as pressure is often ineffective in arousing responses at this period. (3) The head exteroceptors have begun to be functional, and there is evidence that higher brain centers have also begun to be functionally involved in the spinal and brainstem reflexes. This fact is indicated in the alteration in spinal reflexes secured in this late period by cervical transection as reported above. This shift may indeed be associated with an encephalization of function which is in some respects comparable to the encephalization of function which has been treated in other contexts by Fulton and Keller (34), Marquis (60), and others. (4) It is possible that fetal respiration is more effective in the early than in the late fetus.

Stage E. Type II. At varying times, in relation to different receptor areas, as the period for normal birth approaches, the previously recognizable and distinct patterns of response come to be merged in new and larger patterns which are in a novel way effective in changing the orientation of the whole organism, rather than a segment of the organism, in relation to the activating stimulus. This change may, in certain cases, be as much a function of increasing muscular effectiveness as of a central-nervous-system alteration. It is also often apparently a function of the onset of effectiveness in new receptor mechanisms, especially

of a labyrinthine sort. (See Dusser de Barenne, 30). A characteristic example of this type of this stage may be found in the stimulation of a point, such as the nostril, which at first releases neck and fore-limb reactions, but which by 60 days has come to initiate paddling movements which are effective in "avoidance behavior" in which the external observer may see "persistence by varying means" involving the total, integrated organism. Behavior of this sort, if viewed out of the context of the changes just summarized in *Stages A, B, C,* and *D,* is often so complex, variable, and "adaptive" as to seem almost to defy mechanistic description. In the light of the stages just outlined, however, such behavior typically appears as a new temporal pattern of responses which have been present in isolation, under other conditions, at previous stages. The onset of this type of behavior in relation to certain forms of stimulation may occasionally, at any rate, begin quite early in fetal life. In this last stage, then, it seems that a new afferent mechanism may initiate processes leading to old efferent mechanisms and to new time relationships in the operation of these old mechanisms. As a speculation the writer wishes to record the fact that it seems not incredible to him at present that a complete knowledge of this change will make possible an understanding of "equivalence of stimuli" in Klüver's sense (50), and that possibly the processes here briefly characterized will rather explain the conditioned response than be explained by what has customarily been called conditioning.

11. *Photic stimulation of the eye may lead to motor*

responses of the limbs, etc., before the eyes are normally open in the fetus, when such stimulation is made possible by an operative exposure of the eye. In spite of the fact that a microscope was not used in the observations of the eyes, it can be said that the present study demonstrates the existence of the iris reflex as early as the 41st fetal day. Photic stimulation in later fetuses also leads to eyeball movement, eye winking, and the activity of other large muscle groups of the body. Previous observations in this field have been summarized by the author elsewhere (15).

12. *Auditory stimuli released behavior in a 63-day fetus after the liquid had been removed from the external meatus.* No evidence of response to auditory stimuli on the part of fetuses *in situ* or under the surface of the bath was secured, although evidence in comparable forms suggests that under certain circumstances such responses may be secured (15). Avery reports responses in the fetal guinea-pig to auditory stimuli on the 60th day (6).

13. *Needle stimulation ("pain stimuli") in general released responses which were quite comparable to responses released by pressure stimuli rather than to the vigorous responses characteristic of the adult animal when subjected to pain stimulation.* Very little behavioral evidence of "pain," even on actually cutting the skin, was secured in this study. It seems, therefore, that the prepotency of the noci-ceptors, in the sense described by Sherrington, has not developed in any marked degree in fetal life (81). For previous findings which support this view see Genzmer (35).

14. *Temperature stimuli well above and well below the temperature of the fetus, when applied to the skin, release behavior.* The experimentally controlled temperature stimuli upon which this statement is based were applied under such controlled conditions that it could be determined that the responses were to temperature and not to pressure. The sensitivity was first demonstrated in the 35-day fetus. The responses released, possibly because of the spreading of the liquid stimulus, seemed to involve a large number of muscle groups rather than a truly specialized behavior pattern. The possibility that the extremes of temperature used in the present study may have involved pain stimulation rather than true temperature stimulation should be recognized.

15. *In late fetuses compensatory movements during rotation and in the immediate post-rotational period were demonstrated.* In the present study these observations are based upon relatively few experiments. The observations made, however, are clear and it seems possible to make the statement with little qualification. The course of the development of this form of receptor control of behavior during fetal life, however, has not yet been worked out. For an account of previous work in this field see (15).

16. *There is evidence that in late fetal life higher brain centers influence responses which are characteristically called "spinal reflexes."* In a number of fetuses the spinal cord or brain was sectioned at various levels. The results of these preliminary experiments suggest that, in late fetuses at least, the normal spinal reflexes

are in part a function of the activity of higher centers. This finding is in harmony with the facts which have been established in the adult organism by the physiological study of reflex phenomena in the spinal animal (cf. 31, 33, 23). Many of the necessary controls to make assertions sure in regard to the effects of these operations, however, have not as yet been carried out. In general, our somewhat casual results, however, are in accord with those of Brown (10) and Langworthy (54).

17. *As development progresses, the amount of "motor diffusion resulting from specific receptor stimulation" decreases, at least in certain areas and under certain stimulus conditions.* Minkowski, on the basis of his extensive work on human fetal material, in several places speaks of the shrinking of the cutaneous "reflexogenous zones" as age increases (63, 64). Bersot notes a similar change, especially in regard to the responses which are typical of plantar stimulation (8, 9). Pratt, Nelson, and Sun say of the newborn infant that as age increases the child comes to have sucking responses elicited only by the lips which previously were called out by stimulating cheeks, tongue, and the interior of the mouth (75). Pratt, in later publications, suggests that the shrinking of the reflexogenous zones is one of the determinants of specific behavior (71, 72, 73, 74). Before considering this problem in detail, it may be well to define as clearly as possible the concept of the reflexogenous zone. This descriptive phrase has been typically applied to an areal receptor surface of an organism which, when stimulated at any point by a par-

ticular form of energy, releases patterns of motor activity which may be characterized as the same. In practice, the reflexogenous zone must always be determined by appropriately stimulating a given receptor surface and noting the resulting behavior. Thus, for example, when the same stimulus is applied to the upper eyelid and to the concha of the ear, and when the stimulaton of these areas releases, as it sometimes does, an eye wink, these two surfaces, as well as all intermediate and surrounding areas which release this response, may be spoken of as composing the same "reflexogenous zone." This is true even though touching the eyelid does not lead to a pinna response as does touching the concha. Many writers in this field have not clearly separated the concept of the reflexogenous zone, as just defined, from what may be called *motor diffusion to specific receptor stimulation,* when such diffusion is determined by the stimulation of one special receptor area, at various developmental stages. The two concepts, however, of the *reflexogenous zone* and of *motor diffusion to specific receptor stimulation* should be kept distinct. An excellent example of change in a motor diffusion is found in the lateral margin of the pinna (Pinna "A" of the Chart of Receptor Areas given above). At 32 days stimulation of this area led to a patterned movement of neck, trunk, and fore and hind limbs. At 43 days the same sort of stimulation released "a sharp, sudden contraction of the whole pinna, as well as neck, trunk, and limb movement." At 61 days, however, stimulation was followed by "a strictly localized twitch of the exact part

of the pinna that had been touched." This progressive change, according to the definitions offered above, is not an alteration in reflexogenous zone, but in "motor diffusion to specific receptor stimulation," because at all stages the same receptor area was stimulated.

With this distinction in view we may now turn to the question of the shrinking of the zones as development progresses. It has been pointed out that in spite of great variability it may still be asserted that stimulating any specific sensory area in a series of progressive fetal ages characteristically releases what may be considered as a pattern of response. Thus, instead of holding that the "zone" which releases an extension of the digits of the fore paw at 46 days includes the whole cutaneous area of the leg and then in later stages shrinks to the palmar surface of the paw, it may be equally true, and possibly more satisfactory, to consider the characteristic behavior pattern released by *each* of these points at various ages. Thus, when the cutaneous area of the fore leg is considered, digit flexion, as part of a general pattern of limb flexion, may be present at one stage and not at another, but when it is present it may well be considered to be an aspect of the general pattern of limb flexion and not a mere spreading of the reflexogenous zone of the plantar surface. The clinical significance of the responses released by plantar stimulation and the stimuli involved in the release of sucking responses have led *all* acts which include toe movements or mouth movements to be interpreted in relation to the sole of the foot or the lip zone. In the present study different areas were

studied without giving especial weight to any one area. *As a result of this form of investigation*, it seems that many of the alleged changes in reflexogenous zones are seen to be alterations in what has been defined above as motor diffusion to specific receptor stimulation.

The importance of this distinction is made especially clear when the responses released by stimulation of specific mucous-membrane areas are considered, because, as the study of the protocols given above will show, these areas seem to lead to much less variable response, during the changes of fetal life, than do other areas.

The facts of the present study interpreted in the light of the paragraphs just given suggest that the idea of the shrinking of the reflexogenous zone as a function of development is not a universal description of sensory change in fetal life. This "shrinking" is, rather, a description of significance only when it is in the first place clearly defined, and, secondly, when the facts upon which the conclusions, in any given case, are based represent a controlled series of observations and not merely an incidental observation or two.

18. *The present study does not confirm in detail the specific laws of development, alleging that development is in all respects cephalo-caudal, proximo-distal, or from "fundamental" to "accessory" muscles.* There are certain alleged general laws of the development of behavior, of which the most commonly stated are: (*a*) All development begins in the head region and follows temporally a cephalo-caudal course. (*b*) All development of behavior temporally follows a proximo-distal

course "in the free appendages." (*c*) All development of behavior temporally is from the activity of "fundamental" muscles to "accessory" muscles. These three "laws" are primarily concerned with the development of function on the motor side and will, therefore, be dealt with in subsequent papers which are concerned explicitly with the general development of effector action in the fetal guinea-pig.

The observations of the present study present enough unambiguous facts to make clear the reasons for the origin of belief in these generalizations, but so many exceptions to them also are seen to occur that their value as "laws" or as explanatory principles is thrown into question. Mucous-membrane areas, areas with particular sensory function, and those which are especially motile in adult life, as well as certain other special areas, seem to become sensitive first in the developing fetal guinea-pig, as the protocols and summaries given above demonstrate. Certain of these laws have recently been questioned by Pratt (71) and McGraw (61).

19. *It is possible to view most of the typical patterns of behavior released by the stimulation of given areas as capable of securing some end or ends, which during fetal life itself or during subsequent independent life may serve adaptive needs of the organism.* This is not true of all behavior, but it is true of so much behavior that it deserves special treatment here. It should be emphasized that a consideration of *end* or *purpose* in this sense does not, of course, imply what seems to the writer to be the truly preposterous notion

that the purpose of a response, as it is apparent to the external observer, is in any sense concerned in the mechanism which makes this response possible. In spite of the writings of certain modern teleologists of various shades of depravity it still seems scientifically true to assert unequivocally that purpose is functionally unrelated to mechanism. This is, of course, not to say that phylogenetic and ontogenetic processes which have been involved in the development of the mechanisms of the organism in question are not such as to make of the organism an animal which, when placed in a given environment, will respond adaptively. Adaptive responses, occurring at least frequently enough to further the life processes of the organism and thus leading to survival rather than to maladaptation and death, are certainly to be considered as characteristic of every species that survives. Therefore such responses must in a measure be possessed by every surviving individual of a surviving species. As Sherrington says, "In the light of the Darwinian theory every reflex *must* be purposive" (81).

It is possible in this sense to consider most of the typical responses of the fetus as adaptive. This is true whether the behavior released by stimulating a specific area is characterized as a pattern of behavior or as a reflex. It is also possible to view these individual components of behavior as essential aspects or preliminary stages in the adaptive behavior of the total animal when in the late stages of fetal life, or after birth, the organism comes more and more adequately and successfully to carry on, in a world of complex stimuli, the life

processes of a normal and even in a sense "socialized" adult animal.

As the present writer and others have previously pointed out, any classification of teleological activities is one of convenience only (11). Thus, the list of activities given below could be altered according to the system of classification which the classifier might elect to employ. To illustrate this fact an example may be given. Behavior characteristics which are mechanistically unitary may with equal propriety be considered as components of a number of quite diverse adaptive acts. Jaw movement may be thought of as important in both feeding and defense reactions. The outline of purposeful acts given below, therefore, is not intended to be complete or finally valid in any sense, but it is presented merely to illustrate the way in which the mechanistically determined acts of the organism may be classified by cataloguing a series of value judgments made by an external observer. Subject to the qualifications just given, the following classes and examples of behavior acts may be presented in relation to their apparent purpose in the life of the organism.

a. Behavior acts during fetal life involved in the protection, focusing, and use of receptor organs.

1) *Activity of the vibrissae.* A number of investigators have demonstrated the importance of the vibrissae in the adaptive behavior of mammals, and especially of rodents. This evidence has been summarized by Vincent (94). She says that these hairs, due to their generous nerve supply, the leverage by which they magnify tactile stimuli, and the in-

dependent local muscle apparatus by means of which they may be made to vibrate continually, constitute an apparatus which is capable of performing many sensory functions: (*a*) They perform functions as distance receptors because they extend well beyond the body surface. (*b*) They continually sweep the surface before the animal during forward progression and consequently they are important guides in locomotion, and, as Vincent has especially demonstrated, they play a part in the maintenance of equilibrium and in the discrimination of the character of surfaces. (*c*) Failure of stimulation of the vibrissae during forward locomotion elicits violent avoidance reflexes which serve to stop the animal in its forward course. The vibrissae thus warn the animal that it is approaching a possible change of elevation and, as the external observer notes, the animal is thus protected from falling. Concerning the function of the vibrissae in the cat, see Schmidberger (80).

As noted in the protocols, stimulation of the vibrissal areas (2 and 3 of the chart) as early as 32 days leads to neck movements and fore-limb movements which soon come to be of a sort which may be considered as protective; that is, these movements, as such, tend to remove noxious stimuli from this area. It is interesting to note that there is apparently a correlation between the appearance of the vibrissae on the surface of the skin and the onset of behavior released by the stimulation of this area. In a publication appearing since the experiments reported in this paper were done, Harman and Dobrovolny report that vibrissae first ap-

pear on the surface in the 32d day (38). This is exactly the same day when in our experiments responses were first released by stimulating this area. In the course of this activity much self-stimulation of the receptor area itself is effected. Such self-stimulation of the area on each side by the homolateral fore paw has also been observed in the course of this investigation during "spontaneous" behavior at many age levels. At 38 days local corrugations of the skin of the pad begin. The muscle responses which make this corrugation possible are the same which later vibrate the individual vibrissae. From the time that this corrugation appears, stimulation of this region, or of the distal ends of the growing vibrissal hairs, or often of adjoining head regions, usually causes a corrugation of the skin in this sharply localized area. Hence, long before birth any touch stimulus in the head region may often effect an increased vibration and thus bring about what have been called the "tappings of the blindman's cane of the rodent." This response is such that further specific stimulation may be effective in providing a clue which may later be used in the adaptive performance of acts concerned with alimentation, locomotion, reproduction, and the other processes listed below. The twitch of vibrissal hairs to stimulation may be characterized as a local reflex, and yet the study of the development of this reflex given in the protocols presented above shows that this "reflex" was in its inception a component part of the primary behavior pattern involving especially neck and fore-limb movements. Moreover, at any time during fetal life, and indeed after birth, this whole

pattern, and not only the local reflex, can be called out by appropriately stimulating this area. It is apparently also true that, as the so-called encephalization of function becomes increasingly complete in late fetal and early neonatal life, stimulation of this area more and more directs the complex adaptive responses of the whole organism. These responses of the total animal may in some respects be viewed as composed of patterns of response previously typically elicited independently by other forms of stimulation applied to the various specific sensory areas of the body, as tentatively suggested above.

2) *Motor reactions of the visual apparatus.* Evidence has been secured in this study in regard to (*a*) eyeball response, (*b*) iris response, (*c*) eye winking. No evidence has been secured in regard to lens accommodation during the fetal life of the guinea-pig, and indeed the evidence on this matter, as Lashley points out, is not clear even in the adult rodent (55). These motor eye-reactions may be considered in order:

a) *Eyeball movements.* At 35 days the stimulation of the closed lid led to what seemed to be the pulling back of the eyeball into the orbit, as if all of the external eye muscles contracted at once. This was again observed at 38 days. At 40 days, when the closed eyelids were operatively opened, eyeball movement was certainly present. After operative opening of the eyes of fetuses from this time on, eyeball movements were generally observed. After 56 days, when the eyes were first open naturally, eye movement was very frequently seen. Eyeball movements could be elicited by

tactual stimulation of nearby sensory areas, by postural changes, and thus presumably by neck or labyrinthine stimulation, as is typically the case in newborn and adult cats (32, 30). From 45 days on, eyeball movements were observed which seemed to be definitely correlated with photic stimulation of the retina. It seems significant therefore that approximately half way through the gestation period, inside the dark cavity of the mother, with lids still sealed, a neuromuscular mechanism has developed which makes eyeball movement possible. Characteristically at first these movements are local reflexes which are apparently related to more general postural changes. Even when they can be called out independently they still frequently appear as part of a larger, although still in a sense unified, pattern of behavior. It is unnecessary to point out in detail that eyeball movements are significant in a host of adaptive acts in the adult animal which depend upon vision. This is true in the adult guinea-pig to some extent, even though the rodent is usually considered as rather defective in the specificity of visually controlled behavior which it displays (94).

b) *Iris responses.* The guinea-pig is a poor subject for the observation of these responses. The eyes are either so unpigmented in albino animals that the iris is hardly visible, or so deeply pigmented that the whole structure seems to be an even black. In this investigation at 41 and 45 days, however, and at a number of subsequent stages the response of the iris to light was observed. The iris response was seen to occur only to photic stimulation. However, in view of Lashley's find-

ings on the comparable eye of the rat (55), that pupillary change can be elicited by pricking the skin, by shrill sounds, or by general eye movements, it seems possible that the failure to observe such changes in the present study was due to unfavorable conditions and the fact that a microscope was not used in observing the iris when these stimuli were applied. The light stimulus, however, brought about eyeball as well as iris movements, and also other skeletal muscle responses. The significance of this in terms of what is so often referred to as a "pure" example of a fixed reflex—the ciliary reflex—can be gathered by analogy from the development of other specific reflex mechanisms as given in greater detail above.

c) *Eye-wink or palpebral reflex.* Anatomical work on the fetal guinea-pig has shown that the folds of the eyelid arise on the 24th day and fuse to cover the eye on the 29th day (38). Stimulating either the upper or the lower of the fused lids at 32 days led to an elevation of the homolateral limb and soon to head movements as well. Typically the limb was brought up toward the point stimulated. At 35 days a wink, that is, a contraction of what seemed to be the whole orbicularis palpebrarum muscle of the still unopened eye, took place. Head flexion or retraction, elevation and brushing of the stimulated part by the homolateral fore leg, a wink of the still closed lids, therefore, may be characterized as the behavior pattern which is typically released by stimulation of either eyelid. The isolated "palpebral reflex" early comes to be independently elicitable, but until the end of fetal

life the typical pattern involving limb and neck muscles is often called out by directly stimulating either lid. The whole mechanism of this larger pattern may be viewed as protective, as may the eye wink itself. It is interesting to note in this connection that eye-winks, once well established, can at least occasionally be elicited by stimulation of many regions far from the eye itself. Parenthetically, it may be said that this peculiar proneness to release, as shown by certain responses at certain times, is a phenomenon which merits special study during fetal life. The scratch reflex at certain periods seems almost waiting to be set off by a variety of stimuli applied at many different receptor points. This "set-for action" is true also of certain other responses, such as the pinna reflex.

Not only pressure, but photic, thermal, and electrical stimulation, strong mechanical laceration, and body-posture changes provide stimuli which at certain stages elicit the palpebral reflex. In the later part of fetal life, closing of the open eye on a touch of the cornea was an *invariable* response. The biological utility of the eye-wink is obvious and has had many commentators. That its mechanism has been so perfected that it may begin to function at a very early period of active fetal life is significant in evaluating the learning or conditioned-reflex theories of the origin of the eye-wink. It is also significant for the general thesis of this paper that this almost invariable reflex is typically, in early stages and, indeed, under the appropriate conditions at all stages, part of a larger pattern of behavior.

3) *Ear-muscle responses and audition.* Tac-

tual stimulation of the concha or of either margin of the pinna, as well as many peripheral regions, typically elicits what is called the pinna, or Preyer's, reflex. In our records this is noted as appearing on the 36th day. Avery had previously noted it on the 64th day (6). Characteristically this response is part of a much larger pattern of behavior. It will be remembered that the first stimulus-released response recorded in this study was the result of stimulating the concha of the ear. At 31 days such stimulation led to a movement of the fore legs and a bowing of the head. This could be seen as part of the typical pattern released by this type of stimulation. In later stages great variation was noted in the responses elicited by stimulating the ear areas numbered 15, 16, and 17 on the Chart of Receptor Areas given above. In some instances the pinna was elevated, in some it was adducted to the head. In certain cases apparently only the local musculature near the spot stimulated responded, in others the larger muscles at the base of the ear contracted as if to close and protect the concha. Pinna responses of one type or another were occasionally released by touch stimuli at a distance from the ear. Such responses were also elicited at certain ages and under certain conditions by photic stimulation of the eye, by a loud sound, by temperature stimulation, by electrical stimuli, and by needle stimulation. The "purpose" of the pinna response in the adaptive behavior of the adult animal is in a measure obscure. It certainly is typically elicited as part of a general startle response and may possibly have some value as a visual warning signal to other

animals. Certain pinna responses may also assist audition by the raising of what may be called an ear-trumpet to catch vibrations and make localizing responses of the whole organism to auditory stimuli more possible. Pinna and concha responses of a stronger sort apparently also partially close the external meatus and thus, it may be, protect the ear from strong or harmful stimuli. This typical reflex, therefore, like those considered above, is seen to be in its own nature complex and to be characteristically associated with more comprehensive behavior patterns and also possibly to subserve a variety of uses in the adult organism.

4) *Nostril dilation and contraction and olfaction.* Tactual stimulation of the nostril, first at 32 days, released head and fore-paw movements such that, even at this early period, the region stimulated was brushed by the homolateral paw. Later studies showed that it was possible to have the right paw raised by stimulating the right nostril, and vice versa. At 46 days, beside head and trunk reactions, vibrissae pad corrugation, and eye-winks, a definite wrinkling of the muscles of the snout, including those about the nostrils, took place. At 48 days local stimulation led to a sharply localized twitching of the nasalis muscle, together with fore-leg elevation. From this time on this "local reflex" was a typical part of the animal's behavior when appropriately stimulated. Characteristically, however, it remained a component of the more general pattern of response which had preceded it ontogenetically during earlier fetal life. In a fetus delivered into air 3

days before normal birth and kept alive for several weeks, not reported elsewhere in this study, twitchings of the nostrils were recorded which seemed to be associated with olfaction. In this case, however, the whole pattern of breathing was involved. The contraction and dilation of the nostril may thus be considered to be a local reflex which is part of a more general pattern and one which in later life may be considered to be important in regulating olfactory stimuli.

5) *Tongue movements and gustatory stimulation.* Gentle stimulation of the lip by the 37th day led to a slight protrusion of the tongue. Typically, stimulation of the tongue itself, of the lips or gums, led during the early fetal period to head and fore-limb movement, and after the 37th day to tongue movements associated with lip, jaw, and often to other responses of muscles in the head region. In prematurely delivered fetuses it has been observed that the tongue is employed in sucking, and in its movement thus facilitates the general exposure of the sensory surface which is presumably concerned in taste reception. This local tongue response is calculated to play its part in the sensory functions involved in tasting, but it is characteristically also part of a much larger behavior pattern which is significant in the process of eating.

6) *Changes in body posture, the proprioceptors, and the labyrinthine mechanism.* During the latter part of fetal life, changes in body posture were associated with marked tonus changes, including trunk and limb movements. As noted above, some of the Magnus reflexes were apparently elicited by appropriate

stimulation. In air-delivered fetuses, of 65 days, righting reflexes were observed when the animal was placed on its back on a flat surface, as previously noted by Avery at 60 days (6). Compensatory head and limb movements and presumably eye movements occurred on rotation. Post-rotational movements were also observed immediately after rotation ceased. Some of these reactions may have been due to neck-muscle and other forms of proprioceptive and exteroceptive stimulation, but it seems almost certain that a number of these reactions may be attributed to vestibular function. For an analysis of the part played by the various receptors in posture see Dusser de Barenne (30). In the fetal cat vestibular responses have been demonstrated at the 54th day by Fish and Windle (32). In the case of the "Magnus reflexes" some examples of rather localized reactions to this form of stimulation were apparent, but, in general, stimulation of the sort used in this study seems to have involved tonus changes and contractions in associated muscles in much of the whole body. The part that the labyrinthine senses play in the adaptive life of the organism need not be emphasized here. Unpublished protocols suggest that the highly adaptive acts of turning over and of maintaining balance during locomotion are both possible in the prematurely delivered fetus. These responses are, however, still far from perfect in the normal guinea-pig at birth. No observations were made on the air-righting reflex which has been studied in the cat by the present writer (16).

7) *General cutaneous stimulation and motor*

responses. The special touch organs associated with the vibrissae-hairs have been considered above. In almost every pressure stimulus of the many hundreds given in this study the reaction may be thought of as having a bearing upon cutaneous sensitivity, as also may those elicited by deep pressure, temperature, or pain stimuli. Typically, the responses to this form of stimulation are the various behavior patterns noted in detail above. A local response which is frequently part of the general pattern released by cutaneous stimulation is the corrugation of the musculature under the point stimulated. In many cases these responses may be called panniculus carnosus reflexes. Corrugation of this sort moves the hair above the area and makes stimulation more adequate than it would have been otherwise. *Adient,* that is, movements toward, or *abient,* that is, movements away from, in Holt's sense (42), the touch of the nose or the paws, may also be regarded in this manner. Typically, the ability of the organism to localize by local sub-cutaneous muscle contractions or by limb movements the portion of the skin that has been stimulated may be seen to be a function which provides a protective mechanism for removing noxious stimuli. The protocols given above indicate that such stimulation often involves very specific and also quite elaborate patterns of response. These responses are obviously valuable to the organism in removing insect parasites and in effectively dealing with noxious stimuli of various sorts. It is quite possible that this behavior is significant in building up the animal's general ability to deal with stimuli in a spatial framework

("space perception"). This possibility has been considered in some detail in a previous paper by E. T. Raney and the writer (77).

A review of the seven points given above shows that in relation to the principal senses there are certain characteristic "reflexes" concerned in the proper functioning of the receptor mechanisms in question. These reflexes are themselves usually seen to be parts of larger but still specific patterns of response. In many instances these responses have been shown to arise early in motile fetal life and seem too specific and too immediate to have resulted from previous fetal "experience," especially since evidence for such exercise is nowhere available. These patterns of response are possibly best thought of in each case as reactions determined by specific stimulation of a neuromuscular mechanism which is constructed *as it is* at the time. The fine and gross anatomical changes of the neuromuscular mechanisms which result from, or rather are the inner dynamic processes of, morphological growth ("maturation") produce a mechanism which suddenly or gradually becomes capable of entirely new activities when appropriately aroused. In another place the author has considered what some of these mechanistically determined factors of inner growth may be and has attempted to show that they are meaningless if considered as independent of an environment (15). For an important reference in regard to the relationship between the work of Spemann and that of the concept of "physiological gradient" as a determiner of development in the sense considered here, see Child (17).

We may now turn to consider, maintaining our purely pragmatic teleological point of view, how the sense-organ responses just discussed, together with other patterns of behavior, come to play a part in what may be called the major life functions of the organism.

b. Behavior acts concerned in alimentation. During fetal life the elements of the sucking reaction and the biting reaction develop. The patterns involved in these responses include head, lip, gum, jaw, and tongue movements. In most fetal organisms the onset of the feeding reaction is thought of as the onset of sucking, but in the guinea-pig, because of the presence of well-developed teeth in the fetal period, certain of these early movements may be characterized rather as biting than as sucking reactions. Beside these local reactions concerned in feeding, all of the motor mechanisms associated with the special organs of the exteroceptors, all of the mechanisms of locomotion, and indeed, indirectly, almost all of the acts of the organism, may be considered to subserve the essential end of securing, ingesting, and digesting food. If this general view of the development of the feeding reaction be taken, it is obvious that, from the first responses of the fetus, patterns of response are present which will later play a part in this complex adult reaction. The detailed protocols given above indicate the temporal stages at which the special responses of lips, tongue, jaw, head, and trunk first become apparent. The development of the feeding reaction in the guinea-pig is comparable in almost every respect, except the specific age at first appearance, to the development of the sucking reac-

tion in the fetal cat as described by Windle and Griffin (100).

c. Activities of the fetus which are significant in the development of the mechanism of air breathing. Many years ago Ahlfeld pointed out the significance of so-called "premature breathing movements" in organisms whose oxygen supply was still provided by placental circulation (1). The present writer has reviewed something of the history and development of the knowledge of this reaction and the speculation in regard to its function in fetal life (15). In the present study some cases of such movements were observed. Sometimes these movements arose "spontaneously" and sometimes as the result of external stimulation. As an example of the stimulus-aroused movements one may cite the rhythmic chest movements which were set off by a touch of area 36 (the abdomen) in a 41-day fetus. The stimuli which are necessary to initiate breathing movements when placental circulation has been discontinued have been studied (44). The work of Avery indicates that breathing sufficient to provide oxygen for the life processes of the fetus of the guinea-pig appears on the 52nd day (6). If exercise be conceded as advantageous in the preparation of a mechanism for effective function, then the utility of these Ahlfeld movements must be obvious. That they may play a part in fetal life as a sort of "auxiliary pump" to the fetal heart has also been suggested by Walz (95).

d. Responses of the organism concerned with excretion. In the present study stimulation of the areas of the anus (No. 28) and of the external genitals (No.

64) led to a complicated pattern of response which seemed most obviously protective, but which in certain instances definitely involved local and sometimes rhythmic movements of the sphincter muscles. The importance of sphincter control in the adult organism is sufficient evidence of the significance of the onset of this reaction in fetal life. More detailed study like that of Yanase (104, 105) on fetal digestive processes would doubtless show many other fetal activities which could be viewed as preparatory for the function of excretion.

e. Activities of the organism concerned in locomotion. In its most complex form in the adult organism, locomotion may be thought of as involving, in some way, almost all of the motor and sensory mechanisms of the organism and in serving, at least in a subsidiary way, as a means by which most of the organism's important life acts are performed. As Hinsey, Ranson, and McNattin have shown, its central-nervous-system mechanism is also most complex (41). Contrasted to the almost adult form of the guinea-pig at birth is the early fetal condition of the opossum when it is born (40). But even in the opossum the mechanisms of air-breathing, locomotion, and food-taking have developed. Thus, locomotion, like mechanisms for air and food intake, seems to be in most lower mammals a process which must be effective at birth. Without locomotion it is impossible for newborn organisms to find the mother or the protective pouch and thus to secure food and protection, save in the cases where complete maternal care is provided.

Coghill (19) has studied and admirably analyzed the origin of aquatic and terrestrial locomotion in Amblystoma. He has shown the correlation between the growth of locomotion-behavior and the neural mechanisms upon which it depends. Much of this analysis may be suggestively applied to the study of the development of locomotion in other fetal organisms, as Coghill himself has suggested (18). Windle and Griffin have given the results of their own studies and summarized the literature on this subject (100). In an analysis of the onset and development of locomotor activity Coghill's two headings of aquatic and terrestrial locomotion may be used as starting points. During the whole of prenatal life the mammalian fetus is, of course, an aquatic animal. In the fetuses studied in this investigation, long before birth, trunk and limb movements were often performed so that the fetus was propelled through the bath as far as the attachment of the umbilical cord would allow. It can thus be said without doubt that coordinated movements, making effective aquatic locomotion possible, develop well before birth in the fetus of the guinea-pig. In later fetal stages, when the cord was ligated and cut and breathing successfully initiated, terrestrial locomotion was seen to be possible before the period of normal birth. Avery, whose technique emphasized the study of air-breathing fetuses, saw well organized locomotion on the 63rd day (6). In order, however, to consider the developmental onset of locomotion, either aquatic or terrestrial, one must go back to the earliest observed movements of the fetus. Neck flexion and fore-limb

movement appear in our study at 31 days, and both of these responses are involved in locomotion. In the protocols summarized above, reference is made to the onset of independent flexion and extension in the limbs, and to the beginning of the "crossed extension" responses which result from independent stimulation of either fore or hind limbs. Reference is also made to the beginning of "trot" or "paddling" movements, in which the fore limbs alone, the hind limbs alone, or all four limbs together, cooperate in what is sometimes a relatively coordinated form of behavior. Head and trunk flexion and extension are also involved in both swimming and walking. In the guinea-pig fetus the rhythmic activity and the maintained tonus necessary for coordinated swimming are present some time before terrestrial walking becomes possible. The failure of land walking after what seems to be the development of its essential mechanism is probably due to the buoyancy of the organism in liquid. In the naturally-born, full-term guinea-pig it is obvious that the fore legs are more functionally effective than the hind legs. This does not mean that coordinated rhythmic activities are neurologically impossible in the hind limbs at that time, because at a period previous to this they have shown, in organisms in which placental circulation was maintained, and which responded under water, true coordinated "trot" responses. When responding in air, however, the weight of the body which must be supported is too much for the strength of the limb muscles. This observation may be considered as but one example of the fact that much apparent in-

coordination and so-called "mass activity" of the neonatal mammal may be an indication rather of muscular weakness than of an undifferentiated, "unmaturated" nervous system. In conclusion, in regard to locomotion it seems that one can say that swimming and walking in the fetal guinea-pig are dependent upon a mechanism which, structurally, at least in its larger outlines, is ready for action before there is any likelihood that the organism will be in a position to carry on terrestrial locomotion or before swimming would have any obvious biological value for the fetus. It seems to the present writer merely a matter of convenience rather than of fundamental scientific creed whether one elects to view the act of walking as an integration of a series of still partially independent behavior patterns, or whether the subordinate patterns involved in neonatal and fully adaptive locomotion are to be thought of teleologically as a single pattern. In any case, it seems obvious that dozens of detailed responses such as extensor thrusts of isolated limbs, digit flexions, and trunk extension may be separated from this total pattern and intelligibly, if not completely, dealt with as independent entities. That successful land locomotion does not come into being genetically as a complete totality which is in every sense prior to its constituent elementary patterns is, it seems to the writer, demonstrated in the protocols given above and in many others which are not published here.

f. Defense and struggle reactions. These names are used as a traditional, teleological classification under which a large number of the activities of an animal

may be classified. Babinski (7) is, of course, notable as one who has dealt in detail with reactions under this heading. Among other neurologists who have used this means of characterizing behavior may be mentioned Granger and Goltz as reported by Luciani (59), Sherrington (81), Riddoch (79), and Pavlov (68, 69). Such activities as kicking, struggling, biting, pushing away the stimulating object, withdrawing from the stimulating object, turning aside, contracting into something like a ball, sudden muscular quiescence and tonic immobility ("death feigning" or the "still reaction"), shrill vocalization and other expressive movements, exteroceptively initiated locomotion, paw movements, panniculus carnosus responses, the sense-organ protective reflexes noted above, and many other activities, may be included under the classification of defensive acts. Responses over this wide range, it will be obvious, are not unitary, mechanistically describable behavior patterns, but involve in their accomplishment almost every pattern of response of which the organism is capable under some given set of circumstances. Thus, biting is serviceable both as an act of alimentation and as defense; extension of the leg is serviceable in approach and in kicking. The effectiveness of almost all of the patterns of response of the animal in serving in some way a protective need, even during fetal life, is interesting in view of the fact that, as the evidence given in the protocols above shows, the "noci-ceptive system" of Sherrington does not seem especially well developed in fetal life. It is basically impossible to give any general mechanistic description

of the conditions under which activities subserving the end of "defense" are released. These responses, therefore, instead of being considered to possess certain characteristics *because* they are defense reactions, are probably best thought of as a series of mechanistically quite distinct reactions which are only grouped together for the convenience of the observing scientist and are really a result of a series of value decisions on the part of the one who makes the judgments. At certain levels a classification of this sort is, of course, admissible.

g. Sex behavior. During fetal life, none of the specific aspects of sex behavior have made their appearance. It is true, however, that the guinea-pig, in relation to most other mammals, is sexually precocious. During fetal life, moreover, some of the essential general behavioral responses characteristic of the reproductive act have appeared. These reactions include patterns of response in the rump musculature and indeed in the muscles surrounding the external genital organs. The onset of fore-limb movements, trunk responses, and phonation which are concerned in certain patterns in the adult sex act also appear during fetal life. Thus, it can definitely be said that during the prenatal period many responses which will later be constituent elements of coordinated sex acts have made their appearance.

h. Expression and the onset of "emotional" behavior. The somewhat unfortunate term "expression" is still rather generally applied in psychology to those activities of an organism which are considered to be signs of a more general physiological or psychological

state in the animal in question. Often such expressive responses are thought of as important in social stimulation and response and in communication. During normal fetal life one can see the motor onset of various forms of expressive behavior, although little that can yet be truly called expressive, if the social criterion be accepted. The only example of social stimulation which the writer has seen in normal fetal life was in a preparation in which one of two or more quiescent fetuses, still in the unopened uterus, became active and by this activity directly stimulated the contiguous fetus so that it too became active. Essential patterns of emotional expressive behavior, such as baring of teeth, sudden quiescence, erection of hair, and phonation have all been observed in the present investigation during the fetal period. No integrated general pattern of response that can be safely characterized as emotional, however, has been seen in the fetus. But each response which would be involved in a total emotional response of fear, anger, or sex excitement, as described by Darwin (24), and more recently by Dumas (29), Landis (53), and others, is apparent in the fetal period. These responses include limb, paw, body, jaw, lip, nostril, pinna, and pili-erector reactions.

i. Adaptive responses of the total organism in the latter part of fetal life. The various subordinate patterns of adaptive behavior just outlined may sometimes be viewed as components of the total integrated behavioral adaptation of the organism to its environment, as indicated in Type II of *Stage E* of Conclusion 10 above. Some of the components just outlined may be

considered by the external observer who knows what animals do as adults as "preparatory," although obviously when such responses appear in fetal life they are functions of present stimulation and of the structure of the organism as it is at that time.

In fetal life, also, certain complexes of acts already take on a "purposeful" appearance which might be attributed by the untrained or vitalistically biased observer to the "will" or "desire" of the fetus. For example, sometimes after cutaneous stimulation, when a series of brushing movements by the paws has failed to eliminate the stimulus, responses of trunk and limbs take place such that the whole fetus swims rapidly away from the stimulus as far as the attachment of the cord makes mechanically possible. By some observers this action would certainly be characterized as purposeful-avoidance behavior which involved the "docility" (Tolman, 92) of the whole integrated organism and not as a mere series of patterns of responses released by specific stimuli. If the mechanistic analysis of the subordinate patterns of this response, which has been offered above, be remembered, however, it seems not unsatisfactory to consider even this typical total "teleological act" of the integrated organism in terms of an elaborate temporal sequence of stimulus-released behavior acts. These acts, of course, depend on elaborate conditions for their appearance. Work on the facilitation and inhibition of spinal reflexes suggests something of the phenomena which may be found eventually to be basic to alterations of behavior in intact organisms of the sort described. Whatever the

ultimate neural explanation of the apparently "adaptive" and "purposeful" acts of the newborn guinea-pig, it seems sure, however, that these acts need not be considered from all points of view as novel. Nevertheless, if the behavioral criteria of "purpose" be accepted as scientifically satisfactory (which, it may be said parenthetically, as explained above, the present writer does not wish to do), there seems no reason why "purpose" should not be attributed to the total integrated behavior of the fetus during the latter part of the gestation period as well as to the young animal after birth.

20. *The present study does not give unqualified support to any of the more general theories of the development of behavior, such as those summarized by the words "individuation" or "integration," but suggests rather that the formulation of such generalizations is at present premature.* In a recent paper the writer has attempted to review and criticize the theoretical interpretations of the development of behavior with appropriate bibliographical citations (15). From a somewhat different point of view another interpretation of the status of theory in this field has been presented even more recently in an admirable article by Pratt (72). In this paper, therefore, only most brief and general reference will be made to theories of the development of behavior.

If, for the sake of brevity, we omit qualifying phrases, it may be said that two views in regard to the nature of the development of behavior have found supporters in recent years. Two theories of the way in

which this development takes place have also been advocated.

The nature of the development of behavior is held to be either (*a*) the increasing specification of behavior which was previously more general [Cf. Coghill (18), Angulo y González (3), Wheeler and Perkins (97), Ogden and Freeman (66), Irwin (46, 47, 48)], or (*b*) the increasing integration or organization of previously rather discrete elements into new complexes of adaptive behavior [Cf. Watson (96), Pavlov (70), and Minkowski (62) as writers who at least in part advocate this view]. The two theories of the development of behavior which have been sometimes applied to individuation[9] and sometimes to integration are (*a*) the maturation of structures and functions which are fundamental to behavior as a result of hereditary determiners, and (*b*) the environmental conditioning or modification of the structures and functions upon which behavior depends. It would be possible by specific quotation to show that some modern students of the development of behavior emphasize each one of the following four positions: (1) Behavior develops by individuation brought about by maturation. (2) Be-

---

[9]In regard to the use of the terms *individuation* and *specification* by Coghill the following quotation is relevant. "Individuation . . . is not the equivalent of specialization or specification. From the point of view of the organism specialization and specification look outward. They have to do with the adaptation of a part of the organism to a particular function, or relation to the environment. Individuation, on the other hand, looks inward. It is a process by or through which a part of the organism acquires a particular relation to or within the organism itself" (20).

havior develops by integration and organization brought about by maturation. (3) Behavior develops by individuation brought about by environmental action. (4) Behavior develops by integration and organization brought about by environmental action. Intermediate and combination views involving the four positions just given have also been held.

The results of the present investigation, it seems, can not be subsumed under any one of the positions just listed.

A study of the protocols and the conclusions given above shows that the course of development of adaptive behavior, *as it is released by the stimulation of definitely located receptor areas* in a series of increasing temporal stages during fetal life, is in some measure specific for each area studied. Without extending the concepts of integration or individuation until they lack all precise or valuable meaning it does not seem possible to include all of the observations concerning the development of particular behavior systems under any one such heading. For example, the view of individuation as given in the following quotation from Irwin has not always been substantiated in the present investigation, although cases can be found which might be considered to exemplify it.

> . . . patterns differentiate from a primitive general matrix of behavior called mass activity. Mass activity is at its maximum during the first fetal months and during uterine existence the differentiation of activity into patterns is under way (48).

Indeed, in criticism of this view it may be said that

the first ("spontaneous") responses of the organism, which took place before any responses could be elicited by exteroceptive stimulation, were quite specific. These movements involved responses of the muscles of the cephalic end of the trunk and of the fore limb. The first stimulus-aroused activities, likewise, were not "general" but were, on the contrary, quite precise. As fetal development progressed, however, and as other neuro-muscular mechanisms became active, complex responses appeared which, as a whole, might be called "mass-activity." The author has hundreds of feet of motion-picture records of such responses. As noted over and over again in this paper, nevertheless, the responses released by stimulating particular areas at each stage seem to be such that they may be characterized, at least loosely, as special "behavior patterns." In the case of "spontaneous behavior" which appears after many specific patterns have been established, the exciting stimulus is often, it seems, internal to the organism. It is thus conceivable that much of the apparently "random" mass-activity of the fetus (but not the initial responses of the fetus) is a function of internal stimulation which has not yet come to lead to specific responses. Moreover, when activity of so complex a nature that it may be called mass-activity is aroused by experimentally applied exteroceptive stimulation, such activity is usually secondary to the initial release of the more specific behavior pattern ordinarily associated with the receptor area that has been stimulated. An answer may be given that the apparently exteroceptively released "generalized" behavior of the

organism is possibly a function of the proprioceptive or other internal stimulation which itself results from the relatively specific response initially made to the exteroceptive stimulus. Thus, very early, and very late as well, in fetal life, and indeed in neonatal and adult life, strong stimulation or stimulation synchronously affecting a large number of exteroceptive areas or general interoceptive stimulation may produce such complex behavior that it is difficult to analyze it. Even when analyzed with the aid of a motion-picture record such response often seems in fact to justify the term "mass-activity." For an excellent analysis of other meanings of mass-activity see Pratt (72). But controlled exteroceptive stimulation of the sort used in this investigation seems to lead throughout the whole series of fetal stages, in relation to the stimulation of each specific receptor zone, to what may be considered as a special, although very variable, pattern of behavior. These patterns of response, taken as a whole, are at times obscured by other simultaneously aroused activities which are sometimes so general in the number of behavior mechanisms involved that they merit the term "mass-activity."

It is, of course, possible to say that in specific responses the behavior of the whole organism is "implicit," as some of those who hold what is called the "organismic view" assert (47). In a sense such a statement is true. Concerning a response which arises after a previous response, it may be said, for example, that the second response involves "the further extension of the total pattern into the appendages" (19). If the

neck alone flexes, it does always nevertheless bend upon the trunk, and so the trunk may be said to be involved in this neck flexion. In the true "C" reaction, which in many fetuses seems temporally to follow this first neck flexion, the same "totality" of pattern is even more obvious. In many fetuses, however, by the time this reaction is perfected, or even before it has appeared, other special patterns, such as independent forelimb reflexes, have developed. It almost seems that it can be said without too much injustice that such responses may simply be described as they appear, or they may be called "an extension" of something else, depending upon whether one wishes to record an observation or develop a theory. In evaluating the references that have been made to mass-activity, especially by Minkowski and other students, who have observed fetuses that have been removed from placental circulation, the special effects of this abnormal condition should not be overlooked. Progressive asphyxiation in fetal as well as in adult animals not only leads to an alteration of thresholds, but, in certain instances, the accumulation of metabolites so affects the central nervous system that the general response of many muscle groups characteristic of "writhing" appears. Minkowski has discussed this abnormal phenomenon at length (64), and certainly observations made under progressive asphyxiation cannot be compared to observations made under the relatively normal conditions of the present experiment. For a further criticism of the concept of mass-activity, see Dennis (26).

Instead of always beginning in mass-activity, out of

which by *individuation* and *specification* the organism acquires the capacity to use its various appendages and muscle mechanisms independently and adaptively, it seems rather that each specific receptor zone of the organism tends from the first, when stimulated, to release a recognizably specific, but variable, pattern of behavior. Probably ultimately a statistical statement of the exact frequency of released patterns from specific stimulation must be given, as Pratt has done in his recent study of the plantar response (73, 74) and as Gesell and his collaborators (36, 37), Shirley, and many others have done for human neonatal behavior (82).

In the opinion of the author, the fact that random activity cannot be demonstrated to precede all specific behavior acts does not mean that by some sort of mystical hereditary maturation these specific responses unfold to meet some future need of the organism, as the old nativistic theory would have it. Nor are they "instincts" in the old mystical sense. Rather the facts seem to point to the view that, in the complex interactions between the cells and organ systems of the growing individual, factors external as well as internal to the changing systems must be given if a complete story of development is to be presented. As a result of a rigid series of mechanistic determinants which are only now being unraveled by students of developmental mechanics, a receptor-neuro-muscular mechanism which may work more or less effectively when it is first activated as a total arc is produced (cf. 98, 99, 102, 67). In previous papers the author has consid-

ered some of the factors intrinsic and extrinsic to the developing systems themselves which are involved in this process which is often loosely called maturation (12, 13, 14).

It may be said, then, that the facts disclosed in the present investigation do not support the absolute "totalistic" or "organismic" view of the development of behavior. The growth of behavior cannot always be said to involve a simple individuation of previously more general behavior. It is interesting to note in this connection that Minkowski, possibly the greatest student of human fetal behavior, has said in a personal communication to the writer that in his opinion the totalistic approach to behavior is as overemphasized in our time as the atomistic approach was in the last generation.

If we now turn from the extreme organismic view to the extreme atomistic view, we again find a theory that is not supported by the facts recorded in the present study. We are told, for example, that at birth the child "is still a creature of reflexes, although some of these reflexes such as the movements of nursing, occur in fairly predictable series" (85). Such statements illustrate what Koffka has called the impossible effort of Spencer and his followers to build up an understanding of complex adaptive behavior by considering it to involve the integration of specific series of wholly individual reflexes (51). Those who hold the integration-view in its most radical form often suggest that education during prenatal life organizes the adaptive responses which are present in the neonate. In the

present study no evidence has been secured that conditioning or learning occurs so consistently and effectively during fetal life that the adaptive responses of the newborn animal may be viewed as the result of prenatal experience.

In a recent publication Holt says:

> In other words, at this moment (in mammals this is several weeks prior to birth) a pressure on the eye-ball is just as likely to produce (if it produces anything) a twitch of the arm, hand, leg, or toe as it is to produce a movement which in any way refers to the eye. . . . (42)

Since this quotation refers specifically to the stimulation of the eye, the reader may look at the protocols above of the responses given to the stimulation of region 11. Here it certainly seems that from the onset of behavior the responses that appear *are* in some way definitely relevant to the eye. As explained above, it is true, there are possible errors in the experimental observations recorded. At this early stage the observations are based upon comparatively few fetuses. Moreover, it may be suggested that, unknown to the experimenter, a sufficient amount of exercise has been secured to make this apparently patterned reaction, which seems to have been released on truly initial stimulation, really a learned response dependent on previous fetal activity. But this last hypothesis seems unlikely. This is true especially in view of the fact that when the concha alone, of all receptive areas, was sensitive its stimulation released a pattern of response. This early pattern of response, because of mechanical difficulties, such as the shortness of the legs and the like, was not fully effective, but was nevertheless in

outline, as it were, the same movement pattern typical of the adaptive response to be produced in later fetuses from this same sort of stimulation. As suggested in the criticism of the view of mass-activity, the discovery of early specific patterns of behavior, as demonstrated in this investigation, is not to be taken as a ground for a retreat into mysticism or for the apotheosis of a crude preformist view of hereditary maturation. Holt himself, in explaining the education of sensory surfaces, assumes the bony make-up of the jointed limbs as something given *a priori* which is fundamental in determining the nature of first responses. The results of the present study seem to make it necessary, likewise, to assume that ontogenetic factors comparable to those which produce a "patterned" skeleton also produce to some extent, at least, a "patterned" nervous system, ready to work, although often crudely, in a particular way when it is first functionally activated as a total arc. Similar inner structural changes of growth may bring about changes in function at a later period which are not dependent upon previous "total function," at least in all respects. As suggested above, however, what is now known of the developmental mechanics of the embryo indicates that heredity alone will not explain how this mechanism comes to be organized as it is. The observations of fetal behavior made in the course of this study over and over again confirm the self-stimulation of the fetus' sensory surfaces brought about by the fetus' own responses as discussed by Holt (42). Other forms of fetal activity may well play an essential rôle in fetal development, at least at certain periods.

One may therefore still hold, if he wishes to do so, the essential philosophic truth that each individual is at the start of life a *tabula rasa*. This position may be maintained, even though it be understood that natural selection, or some other evolutionary agency, has made adult male and female guinea-pigs capable of producing fertilized germ cells of such chromosomal and cytoplasmic make-up that, *given usual environmental conditions for development,* they will grow into organisms with a series of organized receptor-neuro-motor mechanisms; and that these specific arcs are such that when exteroceptively stimulated for the first time they are capable of producing patterned responses. These first responses, moreover, are at their onset neither in the ordinary sense of the words used "a total reaction of the organism" nor are they "wholly random" in relation to the place stimulated. As development progresses, these responses change. In this process both "individuation" and "integration" are descriptive terms which will characterize particular alterations that may be observed to take place, as has been previously suggested by a number of writers including Minkowski (62), Pavlov (70), Anderson and Patrick (2), and the writer (15).

On the side of positive theory it may be said that the results of the present investigation emphasize the fact that the responses noted at each fetal stage, as they are released by the stimulation of special areas, are most surely correctly described if they be treated as a function of the specific stimulating conditions of the moment and the physiological and anatomical conditions

of the fetal organism as it is at the time of stimulation. Thus, any "long-section view" of a behavior pattern, that is, the description of a behavior pattern through a number of days in fetal life, consists in a notation of modifications in specific responses at a series of temporally separate periods when all possible relevant conditions are known and, if possible, quantified. In many cases, different details of response, or somewhat metaphorically, modifications of behavior patterns, when so recorded, are found in the fetal guinea-pig to follow a temporal course which may, in relation to quite a good many receptor areas, be characterized in terms of the stages considered above under Conclusion 9.

Therefore, in place of asserting that behavior change results from "individuation" or "integration," it seems that the results of the present investigation are best interpreted as a simple record of a series of responses in particular organisms in particular stimulus situations. The scientific purpose of prediction and control of the alterations that may be expected to take place in the behavior patterns of other series of fetal guinea-pigs and, with some reservations, in other series of fetal mammals, may be better served by comparison with the particular characteristics of the series of observations presented here than by any abstract generalizations. Moreover, in spite of two very significant recent theoretical discussions by Stone (89, 90), the present writer still finds the suggested criteria for the radical separation of the influences of heredity and environment, as determiners of the early development

of behavior, impossible of application. For in one sense, at any rate, as the work on developmental mechanics shows, heredity and environment are interdependently involved in the determination even of the very cellular structures upon which all behavior at every level depends (11).

In conclusion, therefore, the author proposes the abandonment of what seem to be premature and all-inclusive generalizations in regard to *integration* or *individuation*, *maturation* or *conditioning*, which are intended to fit the behavioral growth of all responses at all ages in all fetal organisms. In place of such abstractions it seems rather that detailed factual scales, of which the one presented here is an example, should be prepared by those who are interested in early behavioral development. When such descriptive schedules have been recorded for different fetal organisms and for the same fetal organism under experimentally varied circumstances, and when the neurological and other conditions of the observed behavior have been worked out in relation to known stimuli, it may be possible to make valid inter-species comparisons. From such comparisons eventually accurate general laws of behavioral growth may be constructed. Facts which, so far as the author knows, are presented in this study for the first time, moreover, seem to demonstrate clearly that in this typical fetal mammal the development of response, *in relation to specific receptor stimulation* at successive growth stages, is far too specialized to be summarized with validity by any existing general theory of the development of behavior.

IV

SUMMARY

This paper reports an experimental study of the onset and development of reflexes and patterns of behavior in operatively exposed but otherwise normal fetal guinea-pigs ranging from 27 to 67 post-copulation days in age. By the aid of a specially constructed chart, a series of over 100 definite receptor zones were usually stimulated in each fetus. A description of each response, checked by two observers, was recorded in shorthand. Elaborate moving-picture records were also made.

Among others, the following conclusions are based upon the results of this study:

The first true behavior of the fetal guinea-pig involves neck and fore-limb reactions which are released by unknown, i.e., "spontaneous," stimulation at 28 days.

The first exteroceptively aroused reaction was a patterned response, following pressure stimulation of the ear, which involved neck and fore-limb muscle responses at 31 days.

During the course of fetal life, behavior, released by the stimulation of specific receptor areas, may, in spite of great variation at each age, change in a way that may approximately be described as a series of stages. For example, all typical "simple" reflex responses to isolated stimuli can be shown to arise at a given stage which is temporally subsequent to a stage

at which similar stimulation led to a pattern of behavior involving larger muscle groups.

Light, sound, temperature, pain stimuli, non-auditory labyrinthine stimuli, as well as pressure and proprioceptive stimuli, all can be shown to release responses during fetal life.

There is evidence in late fetal life that the higher brain centers influence responses of the sort predominantly mediated at the spinal level.

The concept of the *reflexogenous zone* is distinguished from that of *motor diffusion to specific receptor stimulation,* and the view is advanced that the "shrinking of the reflexogenous zone" is a more limited phenomenon than some writers have supposed.

While strictly maintaining the scientific or "mechanistic" point of view, it is possible for the external observer to recognize many of the responses of the fetus as "adaptive." A number of such responses are considered in detail in regard to their origin and developmental change during fetal life. Among those considered are: (*a*) activity of the vibrissae; (*b*) eyeball movement; (*c*) iris responses; (*d*) eye-winking; (*e*) ear-muscle responses; (*f*) nostril dilation; (*g*) tongue movements; (*h*) changes in neck and body posture in relation to vestibular stimulation; (*i*) general subcutaneous muscle responses; (*j*) behavior acts concerned in feeding, air-breathing, excretion, locomotion, defense, expression, and emotional responses; and (*k*) adaptive responses of the total integrated organisms showing the sort of behavior which has been called "persistence by varied means toward the achievement of a goal" or "docility."

A number of new facts, which so far as the author knows are described for the first time in this study, do not confirm any of the allegedly universal theories of the development of behavior such as those summarized by the concepts of "individuation" or "integration." The specificity of the developmental sequences observed in this study indeed suggests that much further and more accurately quantified work must be done on the growth of behavior in many fetal mammals before the formulation of any such general theory of development may profitably be undertaken.

References

1. Ahlfeld, J. F. Beiträge zur Lehre vom Uebergange der intrauterinen Athmung zur extrauterinen. In *Beiträge zur Physiologie, Festschrift zu Carl Ludwig zu seinem 70. Geburtstage gewidmet von seinen Schülern.* Leipzig: Vogel, 1890. Pp. 1-32.
2. Anderson, A. C., & Patrick, J. R. Some early behavior patterns in the white rat. *Psychol. Rev.,* 1934, 41, 480-496.
3. Angulo y González, A. W. Is myelinogeny an absolute index of behavioral capability? *J. Comp. Neur.,* 1929, 48, 459-464.
4. ———. The prenatal development of behavior in the albino rat. *J. Comp. Neur.,* 1932, 55, 395-442.
5. Avery, G. T. Notes on reproduction in guinea-pigs. *J. Comp. Psychol.,* 1925, 5, 373-396.
6. ———. Responses of foetal guinea pigs prematurely delivered. *Genet. Psychol. Monog.,* 1928, 3, 251-331.
7. Babinski, J. Réflexes de défense. *Brain,* 1922, 45, 149-184.
8. Bersot, H. Variabilité et corrélations organiques. Nouvelle étude du réflexe plantaire. *Schweiz. Arch. f. Neur. u. Psychiat.,* 1918, 4, 277-323; 1919, 5, 305-324.
9. ———. Développement réactionnel et réflexe plantaire du bébé né avant terme à celui de deux ans. *Schweiz. Arch. f. Neur. u. Psychiat.,* 1920, 7, 212-239; 1921, 8, 47-74.
10. Brown, T. G. On the activities of the central nervous system of the un-born foetus of the cat; with a discussion of the question whether progression (walking, etc.) is a "learnt" complex. *J. Physiol.,* 1914-1915, 49, 208-215.
11. Carmichael, L. Heredity and environment: are they antithetical? *J. Abn. & Soc. Psychol.,* 1925, 20, 245-260.
12. ———. The development of behavior in vertebrates experimentally removed from the influence of external stimulation. *Psychol. Rev.,* 1926, 33, 51-58.
13. ———. A further study of the development of behavior in vertebrates experimentally removed from the influence of external stimulation. *Psychol. Rev.,* 1927, 34, 34-47.
14. ———. A further experimental study of the development of behavior. *Psychol. Rev.,* 1928, 35, 253-260.

15. ———. Origin and prenatal growth of behavior. In *A handbook of child psychology*, 2nd ed. rev., ed. by C. Murchison. Worcester, Mass.: Clark Univ. Press; London: Oxford Univ. Press, 1933. Pp. 31-159.

16. ———. The genetic development of the kitten's capacity to right itself in the air when falling. *J. Genet. Psychol.*, 1934, 44, 453-458.

17. CHILD, C. M. *Physiological dominance and physiological isolation in development and reconstitution. Arch. f. Entwicklungsmechanik der Organismen*, 1929, 117, 21-66.

18. COGHILL, G. E. The early development of behavior in Amblystoma and in man. *Arch. Neur. & Psychiat.*, 1929, 21, 989-1009.

19. ———. Anatomy and the problem of behaviour. Cambridge, England: Univ. Press; New York: Macmillan, 1929. Pp. xi+113.

20. ———. Correlated anatomical and physiological studies of the growth of the nervous system of Amphibia. XI. The proliferation of cells in the spinal cord as a factor in the individuation of reflexes of the hind leg of *Amblystoma punctatum*, Cope. *J. Comp. Neur.*, 1933, 57, 327-358.

21. COBB, D. The examination of the central nervous system. Edinburgh: Livingstone, 1928. Pp. xi+248.

22. CORONIOS, J. D. Development of behavior in the fetal cat. *Genet. Psychol. Monog.*, 1933, 14, 283-286.

23. CREED, R. S., DENNY-BROWN, D., ECCLES, J. C., LIDDELL, E. G. T., & SHERRINGTON, C. S. Reflex activity of the spinal cord. Oxford: Clarendon Press, 1932. Pp. vi+183.

24. DARWIN, C. R. Expression of emotion in man and animals. New York: Appleton, 1873. Pp. vi+374.

25. DEMPSEY, E. W., MYERS, H. I., YOUNG, W. C., & JENNISON, D. B. Absence of light and the reproductive cycle in the guinea-pig. *Amer. J. Physiol.*, 1934, 109, 307-311.

26. DENNIS, W. The rôle of mass activity in the development of infant behavior. *Psychol. Rev.*, 1932, 39, 593-595.

27. ———. A description and classification of the responses of the newborn infant. *Psychol. Bull.*, 1934, 31, 5-22.

28. DRAPER, R. L. The prenatal growth of the gunea-pig. *Anat. Rec.*, 1920, 18, 369-392.

29. Dumas, G. Nouveau traité de psychologie. Tome deuxième. Paris: Librarie Félix Alcan, 1932. Pp. vi+612.

30. Dusser de Barenne, J. G. The labyrinthine and postural mechanisms. In *A handbook of general experimental psychology*, ed. by C. Murchison. Worcester, Mass.: Clark Univ. Press; London: Oxford Univ. Press, 1934. Pp. 204-246.

31. Fearing, F. Reflex action; a study in the history of physiological psychology. Baltimore: Williams & Wilkins, 1930. Pp. xiii+350.

32. Fish, M. W., & Windle, W. F. The effect of rotatory stimulation on the movements of the head and eyes in newborn and young kittens. *J. Comp. Neur.*, 1932, 54, 103-107.

33. Fulton, J. F. Muscular contraction and the reflex control of movement. Baltimore: Williams & Wilkins, 1926. Pp. xv+644.

34. Fulton, J. F., & Keller, A. D. The sign of Babinski, a study of the evolution of cortical dominance in primates. Springfield, Ill.: Thomas, 1932. Pp. xi+165.

35. Genzmer, A. Untersuchungen über die Sinneswahrnehmungen des neugeborenen Menschen. (Dissertation, 1873.) Halle: Niemeyer, 1882. Pp. 28.

36. Gesell, A. Infancy and human growth. New York: Macmillan, 1928. Pp. xvii+418.

37. Gesell, A., & Thompson, H. Infant behavior: Its genesis and growth. New York: McGraw-Hill, 1934. Pp. viii+343.

38. Harman, M. T., & Dobrovolny, M. P. The development of the external form of the guinea pig (Cavia cobaya) between the ages of 21 days and 35 days of gestation. *J. Morph.*, 1933, 54, 493-515.

39. Harman, M. T., & Prickett, M. The development of the external form of the guinea pig (Cavia cobaya) between the ages of eleven days and twenty days of gestation. *Amer. J. Anat.*, 1932, 49, 351-378.

40. Hartman, C. G. Studies in the development of the opossum Didelphys Virginiana L. V. The phenomena of parturition. *Anat. Rec.*, 1920, 19, 251-261.

41. Hinsey, J. C., Ranson, S. W., McNattin, R. F. The rôle of the hypothalamus and mesencephalon in locomotion. *Arch. Neur., & Psychiat.*, 1930, 23, 1-42.

42. HOLT, E. B. Animal drive and the learning process. Vol. I. New York: Holt, 1931. Pp. xi+307.

43. HOOKER, D. The development and function of voluntary and cardiac muscle in embryos without nerves. *J. Exper. Zool.*, 1911, 11, 159-186.

44. HUGGET, A. ST. G. Foetal respiratory reflexes. *J. Physiol.*, 1930, 69, 144-152.

45. IBSEN, H. L. Prenatal growth in guinea-pigs with special reference to environmental factors affecting weight at birth. *J. Exper. Zool.*, 1928, 51, 51-91.

46. IRWIN, O. C. The organismic hypothesis and differentiation of behavior. I. The cell theory and the neurone doctrine. *Psychol. Rev.*, 1932, 39, 128-146.

47. ———. The organismic hypothesis and differentiation of behavior. II. The reflex arc concept. *Psychol. Rev.*, 1932, 39, 189-202.

48. ———. The organismic hypothesis and differentiation of behavior. III. The differentiation of human behavior. *Psychol. Rev.*, 1932, 39, 387-393.

49. ISHII, O. Observations on the sexual cycle of the guinea-pig. *Biol. Bull.*, 1920, 38, 237-250.

50. KLÜVER, H. Behavior mechanisms in monkeys. Chicago: Univ. Chicago Press, 1933. Pp. xvii+387.

51. KOFFKA, K. The growth of the mind: An introduction to child-psychology. (Trans. by R. M. Ogden.) 2nd ed., rev. New York: Harcourt, Brace; London: Kegan Paul, 1928. Pp. xix+427.

52. KUO, Z. Y. Ontogeny of embryonic behavior in Aves. II. The mechanical factors in the various stages leading to hatching. *J. Exper. Zool.*, 1932, 62, 453-489.

53. LANDIS, C. Emotion: II. The expressions of emotion. In *A handbook of general experimental psychology*, ed. by C. Murchison. Worcester, Mass.: Clark Univ. Press; London: Oxford Univ. Press, 1934. Pp. 312-351.

54. LANGWORTHY, O. R. A correlated study of the development of reflex activity in fetal and young kittens and the myelinization of tracts in the nervous system. *Carnegie Instit. Wash., Contrib. Embryol.*, 1929, 20, No. 114, 127-171.

55. LASHLEY, K. S. The mechanism of vision. V. The structure and image-forming power of the rat's eye. *J. Comp. Psychol.*, 1932, 13, 173-200.

56. Loeb, L. The cyclic changes in the ovary of the guinea-pig. *J. Morph.*, 1911, 22, 37-70.

57. ———. The correlation between the cyclic changes in the uterus and the ovaries in the guinea-pig. *Biol. Bull.*, 1914, 27, 1-44.

58. ———. Corpus luteum and the periodicity in the sexual cycle. *Science*, N. S., 1918, 48, 273-277.

59. Luciani, L. *Human physiology*, Vol. III. (*Trans. by F. A. Welby.*) London: Macmillan, 1915. Pp. x+667.

60. Marquis, D. G. Effects of removal of the visual cortex in mammals, with observations on the retention of light discrimination in dogs. *Proc. Asso. Res. in Nervous & Mental Disease*, 1932, 13, 558-592.

61. McGraw, M. B. Discussion: Grasping in infants and the proximo-distal course of growth. *Psychol. Rev.*, 1933, 40, 301-302.

62. Minkowski, M. Zum gegenwartigen Stand der Lehre von den Reflexen in entwicklungsgeschichtlicher und der anatomisch-physiologischer Beziehung. *Schweiz. Arch. f. Neur. u. Psychiat.*, 1924, 15, 239-259; 1925, 16, 133-152, 266-284.

63. ———. Sur les modalités et la localisation du réflexe plantaire au cours de son évolution du foetus à l'adulte. *C. r. du Cong. des méd. aliénistes et neurologistes fr.*, Geneva, 1926, 30, 301-308.

64. ———. Neurobiologische Studien am menschlichen Foetus. *Abderhaldens Hdbh. d. biol. Arbeitsmeth.*, 1928, Abt. V, T. 5B, H. 5, 511-618.

65. ———. Über die elektrische Erregbarkeit der fötalen Muskulatur. *Schweiz. Arch. f. Neur. u. Psychiat.*, 1928, 22, 64-71.

66. Ogden, R. M., & Freeman, F. S. Psychology and education. Rev. ed. New York: Harcourt, Brace, 1932. Pp. xiv+350.

67. Orr, D. W., & Windle, W. F. The development of behavior in chick embryos: the appearance of somatic movements. *J. Comp. Neur.*, 1934, 60, 271-283.

68. Pavlov, I. P. Conditioned reflexes: An investigation of the physiological activity of the cerebral cortex. (Trans. by G. V. Anrep.) London: Oxford Univ. Press, 1927. Pp. xv+430.

69. ———. Lectures on conditioned reflexes: Twenty-five years of objective study of the higher nervous activity (behaviour) of animals. (Trans. by W. H. Gantt with the collaboration of G. Volborth.) New York: International Publishers, 1928. Pp. 414.

70. ———. The reply of a physiologist to psychologists. *Psychol. Rev.*, 1932, 39, 91-127.

71. PRATT, K. C. The neonate. In *A handbook of child psychology*, 2nd ed. rev., ed. by C. Murchison. Worcester, Mass.: Clark Univ. Press; London: Oxford Univ. Press, 1933. Pp. 163-208.

72. ———. Specificity and generalization of behavior in newborn infants: a critique. *Psychol. Rev.*, 1934, 41, 265-284.

73. ———. Generalization and specificity of the plantar response in newborn infants. The reflexogenous zone: I. Differential sensitivity and effector-segment participation according to the area of stimulation. *J. Genet. Psychol.*, 1934, 44, 265-300.

74. ———. Generalization and specificity of the plantar response in newborn infants. The reflexogenous zone: II. Segmental patterning of responses. *J. Genet. Psychol.*, 1934, 45, 22-38.

75. PRATT, K. C., NELSON, A. K., & SUN, K. H. The behavior of the newborn infant. *Ohio State Univ. Stud.: Contrib. Psychol.*, 1930, No. 10. Pp. ix+237.

76. PREYER, W. Specielle Physiologie des Embryo. Untersuchungen ueber die Lebenserscheinungen vor der Geburt. Leipzig: Grieben, 1885. Pp. 644.

77. RANEY, E. T., & CARMICHAEL, L. Localizing responses to tactual stimuli in the fetal rat in relation to the psychological problem of space perception. *J. Genet. Psychol.*, 1934, 45, 3-21.

78. READ, J. M. The intra-utrine growth-cycles of the guinea-pig. *Arch. f. Entwicklungsmechanik der Organismen*, 1913, 35, 708-723.

79. RIDDOCH, G. The reflex functions of the completely divided spinal cord in man, compared with those associated with less severe lesions. *Brain*, 1917, 40, 264-402.

80. SCHMIDBERGER, G. Über die Bedeutung der Schnurrhaare bei Katzen. *Zsch. f. vergl. Physiol.*, 1932, 17, 387-407.

81. SHERRINGTON, C. S. The integrative action of the nervous system. New Haven, Conn.: Yale Univ. Press, 1906. Pp. xvi+411.

82. SHIRLEY, M. M. The first two years: a study of twenty-five babies. (3 vols.) Minneapolis, Minn.: Univ. Minnesota Press, 1931-1933. Pp. 221; 513; 228.

83. SKINNER, B. F. The concept of the reflex in the description of behavior. *J. Gen. Psychol.*, 1931, 5, 427-458.

84. ———. The measurement of "spontaneous activity." *J. Gen. Psychol.*, 1933, 9, 3-24.

85. SMITH, S., & GUTHRIE, E. R. General psychology in terms of behavior. New York: Appleton, 1921. Pp. xii+270.

86. STOCKARD, C. R., & PAPANICOLAOU, G. N. The existence of a typical oestrous cycle in the guinea-pig—with a study of its histological and physiological changes. *Amer. J. Anat.*, 1917, 22, 225-265.

87. ———. The existence of a typical oestrous cycle in guinea-pigs and its histology. *Anat. Rec.*, 1917, 11, 411-412.

88. ———. A rhythmical "heat period" in the guinea-pig. *Science*, N. S., 1917, 46, 42-44.

89. STONE, C. P. Learning: I. The factor of maturation. In *A handbook of general experimental psychology*, ed. by C. Murchison. Worcester, Mass.: Clark Univ. Press; London: Oxford Univ. Press, 1934. Pp. 352-381.

90. ———. Maturation and "instinctive" functions. In *Comparative psychology*, ed. by F. A. Moss. New York: Prentice-Hall, 1934. Pp. 37-72.

91. SWENSON, E. A. Motion pictures of activities of living albino-rat fetuses. *Anat. Rec.*, 1928, 38, 63.

92. TOLMAN, E. C. Purposive behavior in animals and men. New York: Century, 1932. Pp. xiv+463.

93. TRACY, H. C. The development of motility and behavior reactions in the toadfish (Opsanus Tau). *J. Comp. Neur.*, 1926, 40, 253-360.

94. VINCENT, S. B. The function of the vibrissae in the behavior of the white rat. *Behav. Monog.*, 1912, 1, No. 5. Pp. 81.

95. WALZ, W. Über de Bedeutung der intrautrinen Atembewegungen. *Monatsschr. f. Geburtshilfe u. Gynäkol.*, 1922, 60, 331-341.

96. WATSON, J. B. The psychological care of infant and child. New York: Norton, 1928. Pp. 195.

97. WHEELER, R. H., & PERKINS, F. Principles of mental development. New York: Crowell, 1932. Pp. xxvi+529.

98. WINDLE, W. F. Correlation between the development of local reflexes and reflex arcs in the spinal cord of cat embryos. *J. Comp. Neur.*, 1934, 59, 481-505.

99. WINDLE, W. F., FISH, M. W., & O'DONNELL, J. E. Myelogeny of the cat as related to development of fiber tracts and prenatal behavior patterns. *J. Comp. Neur.*, 1934, 59, 139-157.

100. WINDLE, W. F., & GRIFFIN, A. M. Observations on embryonic and fetal movements of the cat. *J. Comp. Neur.*, 1931, 52, 149-188.

101. WINDLE, W. F., O'DONNELL, J. E., & GLASSHAGLE, E. E. The early development of spontaneous and reflex behavior in cat embryos and fetuses. *Physiol. Zool.*, 1933, 6, 521-541.

102. WINDLE, W. F., & ORR, D. W. The development of behavior in chick embryos: spinal cord structure correlated with early somatic motility. *J. Comp. Neur.*, 1934, 60, 287-306.

103. WINDLE, W. F., ORR, D. W., & MINEAR, W. L. The origin and development of reflexes in the cat during the third fetal week. *Physiol. Zool.*, 1934, 7, 600-617.

104. YANASE, J. Beiträge zu Physiologie der persistaltischen Bewegungen des embryonalen Darmes. I. Mitteilung. *Pflüg. Arch. f. d. ges. Physiol.*, 1907, 117, 345-383.

105. ———. Beiträge zu Physiologie der persistaltischen Bewegungen des embryonalen Darmes. II. Mitteilung. Beobachtungen an menschlichen Föten. *Pflüg. Arch. f. d. ges. Physiol.*, 1907, 119, 451-464.

106. YOUNG, W. C., MYERS, H. I., & DEMPSEY, E. W. Some data from a correlated anatomical, physiological, and behavioristic study of the reproductive cycle in the female guinea pig. *Amer. J. Physiol.*, 1933, 105, 393-398.

UNE ÉTUDE EXPÉRIMENTALE CHEZ LE COCHON D'INDE FOETAL DE L'ORIGINE ET DU DÉVELOPPEMENT DES RÉFLEXES ET DES FORMES DU COMPORTEMENT PAR RAPPORT À LA STIMULATION DES AIRES RÉCEPTRICES SPÉCIFIQUES PENDANT LA PÉRIODE DE LA VIE FOETALE ACTIVE

(Résumé)

Cet article rapporte une étude expérimentale du commencement et du *développement des réflexes et des formes du comportement* chez des cochons d'Inde foetaux opérativement exposés mais autrement normaux âgés de 27 à 67 jours après la copulation. On a le plus souvent stimulé une série de plus de 100 zones réceptrices définies chez chaque foetus. On a noté sténographiquement une description de chaque réponse, contrôlée, par deux *observateurs. On a fait aussi des notations cinématographiques élaborées.*

Le premier vrai comportement du cochon d'Inde foetal comprend des réactions du cou et des avant-membres lesquelles sont déchargées par une stimulation inconnue, c'est-à-dire, "spontanée," à 28 jours.

La première réaction extéroceptivement déchargée a été une réponse d'une forme définie, après une stimulation de l'oreille par la pression, qui a compris des réponses musculaires du cou et des avant-membres à 31 jours.

On peut montrer que toutes les réponses réflexes typiques aux stimuli isolés *se montrent à une stage donnée qui est subséquente à l'égard du temps* à une stage où une stimulation pareille cède une forme de comportement qui comprend de plus grands groupes musculaires.

On montre que les stimuli de lumière, de son, de température, les stimuli de douleur, les stimuli labyrinthiques non auditifs, ainsi que les stimuli de pression et les stimuli proprioceptifs, déchargent tous des réponses pendant la vie foetale.

Il y a de l'évidence dans la vie foetale avancée que les centres cérébraux supérieurs influençent les réflexes de la moelle épinière.

On distingue le concept de la *zone réflexogène* de celui de la *diffusion motrice à la stimulation réceptrice spécifique,* et l'on avance l'idée que "la diminution de la zone réflexogène" est un phénomène plus limité que quelques auteurs l'ont supposé.

On considère nombre de réponses adaptives en détail à l'égard de leur origine et de leur changement du développement pendant la vie foetale. Parmi celles considérées sont: *(a)* l'activité des vibrisses; *(b)* le mouvement du globe de l'oeil; *(c)* les réponses de l'iris; *(d)* les clignements d'yeux; *(e)* les réponses des muscles des oreilles; *(f)* la dilation des narines; *(g)* les mouvements de la langue; *(h)* les changements de la posture du cou et du corps par rapport à la stimulation vestibulaire; *(i)* les réponses générales des muscles sous-cutanés; *(j)* les actes du comportement dont il s'agit dans les réponses de prise de nourriture, de respiration de l'air, d'excrétion, de locomotion, de défense, d'expression et les réponses émotives; *et (k) les réponses adaptives de tout l'organisme intégré.*

Les résultats de cette étude ne confirment en détail aucunes des théories affirmées universelles du développement du comportement telles que celles résumées par les concepts de "l'individuation" ou de "l'intégration."

CARMICHAEL

EINE EXPERIMENTELLE UNTERSUCHUNG BEIM MEERSCHWEINCHEN VOR DER GEBURT ÜBER DIE ENTWICKLUNG DER REFLEXE UND GEBILDE DES VERHALTENS IM VERHÄLTNIS ZU DER ANREIZUNG DER SPEZIFISCHEN EMPFÄNGERFLÄCHEN WÄHREND DER ZEIT DES AKTIVEN FÖTUSLEBENS

(Referat)

Diese Abhandlung berichtet über eine experimentelle Untersuchung des Anfangs und der Entwicklung der Reflexe und Gebilde des Verhaltens bei operativ ausgesetzten, aber sonst normalen Fötusmeerschweinchen im Alter von 27 bis 67 Tagen nach der Paarung. Eine Reihe von über 100 bestimmten Empfängerzonen wurden gewöhnlich bei jeden Fötus angereizt. Eine Beschreibung von jeder Reaktion wurde in Kurzschrift von zwei Beobachtern aufgeschrieben. Sorgfältige photographische Aufnahmen wurden auch gemacht.

Die ersten wahren Verhaltungsweisen des Fötusmeerschweinchens bestehen aus Hals- und Vorderbeinreaktionen, die durch unbekannte, d.h. "freiwillige" Anregung zu 28 Tagen ausgelöst werden.

Die erste exterozeptiv ausgelöste Reaktion war eine Musterreaktion, die auf Druckreizung des Ohres folgte und die auch Hals- und Vorderbeinmuskelreaktionen zu 31 Tagen einschloss.

Alle typischen Reflexreaktionen auf isolierte Reize können bei einem gegebenen Stadium nachgewiesen werden, die zeitlich einem Stadium folgt, bei welchem ähnliche Reizung zu einem Gebilde des Verhaltens führt, das aus grösseren Muskelgruppen besteht.

Licht-, Laut-, Temperatur-, Schmerzreize, nichtlabyrinthische Gehörreize, sowie Druck- und Propriozeptivreize erzeugen Reaktionen während des Fötuslebens, wie nachgewiesen wurde.

Es gibt Evidenz beim späten Fötusleben, dass die höheren Hirnzentren die Rückenmarkreflexe beeinflussen.

Der Begriff der reflexogenen Zonen unterscheidet sich von der der *Motorausbreitung auf spezifische Empfängerreizung*, und die Ansicht wird vorgeschlagen, dass das "Schrumpfen der reflexogenen Zone" ein beschränkteres Phänomen ist, als einige Autoren vermutet haben.

Eine Anzahl Anpassungsreaktionen wird in Einzelheiten in bezug auf ihren Ursprung und die Entwicklungsveränderung während des Fötuslebens betrachtet. Unter den betrachteten Einzelheiten befinden sich: *(a)* Tätigkeit der vibrissae; *(b)* Augapfelbewegung; *(c)* Irisreaktionen; *(d)* Blinzeln mit den Augen; *(e)* Ohrenmuskelreaktionen; *(f)* Nasenlochausdehnung; *(g)* Zungenbewegungen; *(h)* Veränderungen der Hals- und Körperhaltung im Verhältnis zur Vorhofreizung; *(i)* allgemeine Muskelreize unter der Haut, *(j)* Verhaltungsweisen in bezug auf Essen, Atmen, Ausscheidung, Ortsveränderung, Verteidigung, Ausdrucks- und Gemütsreakionen, und *(k)* Anpassungsreaktionen des ganzen integrierten Organismus.

Die Befunde dieses Studiums bestätigen durchaus nicht die angeblich universellen Lehren der Entwicklung des Verhaltens wie jene, die durch die Begriffe der "Individuation" oder "Integration" zusammengefasst werden.

CARMICHAEL